

Algoritmo de asignación escolar - Perú 2022

1. Descripción de inputs

A continuación se describen las tablas que contienen la información sobre vacantes y postulantes necesarias para llevar a cabo el proceso de asignación escolar.

Vacancies: Esta tabla contiene información sobre cada uno de los programas ofrecidos por las instituciones. Cada institución especifica su disponibilidad para cada uno de los grados de ofrece, incluyendo vacantes regulares y vacantes para estudiantes con Necesidades Educativas Especiales (NEE). A continuación se describen los campos requeridos de esta tabla. Otros campos con información adicional pueden ser añadidos.

- **localId:** especifica el código local de la institución
- **servicId:** especifica el código de la institución
- **gradId:** grado para el que se ofrecen las vacantes
- **shiftId:** turno para el que se ofrecen las vacantes
- **totalVacancyNna:** cantidad de vacantes regulares ofrecidas
- **totalVacancyNnaNee:** cantidad de vacantes ofrecidas para estudiantes con NEE
- **levelId**
- **studentBodyId**
- **studentModalityId**
- **classroomTypeId**

Postulations: Esta tabla relaciona a cada postulante con su representante legal. Además se especifica el tipo de postulación que se está realizando (individual o grupal). Los campos de interés en esta tabla son:

- **postulantId:** Identificador único del estudiante
- **guardianId:** Identificador único del representante. Un mismo representante puede tener asociado varios estudiantes, como el caso en que se tiene hermanos.
- **typId:** este campo contiene una "I" si la aplicación se hace individual, y una "G" si la aplicación se hace de manera grupal.

Postulants: Contiene información sobre los estudiantes, como su nombre, ubicación, representante legal y si el estudiante posee NEE. Los campos de interés para esta tabla son:

- **postulantId:** identificador único del estudiante
- **priority:** es un booleano que contiene "True" en los casos en que los estudiantes poseen NEE.

Guardians: Contiene información personal sobre los representantes legales de los estudiantes. Esta tabla es de carácter informativo y ninguno de sus campos interviene en el proceso de asignación de cupos.

Demand: Esta tabla especifica los programas a los que está aplicando cada uno de los estudiantes. Contiene los siguientes campos:

- **postulantId:** identificador único del estudiante
- **gradeld:** grado al que está aplicando el estudiante
- **order:** un número entre 1 y n que indica el orden de selección de esa institución dentro de la postulación del estudiante.
- **serviceld:** institución educativa a la que el estudiante aplica

2. Descripción de las tablas que consume el algoritmo

Actualmente, el algoritmo de asignación diferida está desarrollado para extraer la información desde tablas con una estructura diferente a la mencionada anteriormente. Por lo tanto, es necesario llevar a cabo un procesamiento y transformación de datos para llevarlos a un formato que pueda ser consumido por el algoritmo. A continuación se describe la estructura de los datos necesarios para ejecutarlo con éxito.

Vacancies: Archivo indicando los programas que serán parte del DA (algoritmo de asignación diferida). Esta base debe incluir las siguientes columnas:

- **Program_id:** Identificador de un programa. Toma valores float o str
- **Quota_id:** Identificador de cuota
- **Institution_id:** Identificador de la institución. Genera relaciones entre varios programas.
- **Grade_id:** Grado (o nivel) del programa.
- **Regular_vacancies:** Cantidad de vacantes de tipo regular.
- **special_i_vacancies:** (Opcional) Cantidad de vacantes de tipo especial. Se espera que i tome valores 1 hasta n, con n la cantidad de vacantes de tipo especial.

Applicants: Archivo indicando a los postulantes al DA. Esta base debe incluir las siguientes columnas:

- **Applicant_id:** Identificador de postulante. Toma valores float o str.
- **Grade_id:** Grado (o nivel) al que postula.
- **Special_assignment:** Int de 0 a n indicando el tipo de asignación, 0 es regular.
- **Secured_enrollment_program_id:** (Opcional) En caso de poseer secured enrollment, corresponde a un id de programa. En caso contrario puede tomar valores None o 0.
- **Secured_enrollment_quota_id:** (Opcional) En caso de poseer secured enrollment, corresponde a un id de cuota. En caso contrario puede tomar valores None o 0.
- **Applicant_characteristics_i:** (Opcional) Columnas con características relevantes para alterar el orden de postulación a cuotas. Se espera que i tome valores 1 hasta n, con n la cantidad de características.

Applications: Archivo indicando las postulaciones de cada estudiante y la prioridad y lotería de estas. Esta base debe incluir las siguientes columnas:

- **Applicant_id:** Corresponde a un id de postulante.
- **Program_id:** Corresponde a un id de programa.

- **Quota_id:** Corresponde a un id de cuota.
- **Institution_id:** Corresponde al id de institución correspondiente al programa.
- **Ranking_program:** Lugar de el programa en la postulación del postulante. Toma valores 0 a n con n la cantidad de postulaciones del postulante.
- **Priority_profile_program:** Int. Corresponde a un de los perfiles de prioridad del la asignación.
- **Lottery_number_quota:** Float. Número de lotería correspondiente al programa y cuota de la postulación.
- **priority_number_quota:** Float. Número de prioridad correspondiente al programa y cuota de la postulación. Esta determinado por el perfil de prioridad.

Priority_profiles: Archivo indicando los perfiles de prioridad, indicando además las transiciones de números de prioridad para el caso de sibling dinámico.

- **Priority_profile:** Int. Número de prioridad. Se espera tome valores 1 a n, con n la cantidad de perfiles de prioridad.
- **Priority_qi:** Prioridad de la cuota i. Se espera que i tome valores según la cantidad de cuotas presentes.
- **Priority_profile_sibling_transition:** Int. Transición del perfil de prioridad en caso de recibir sibling_priority. Corresponde a uno de los perfiles de prioridad.

Quota_order: Archivo indicando el orden de postulación a las cuotas que le corresponde a distintas combinaciones de perfil de prioridad y características del postulante que no postulan simplemente 1-2-3-4.

- **Priority_profile:** Int. Corresponde a uno de los perfiles de prioridad.
- **Secured_enrollment_indicator:** Bool. True si el reordenamiento corresponde a un priority profile de secured enrollment.
- **Secured_enrollment_quota_id_criteria:** Criterio a cumplir para aplicar el reordenamiento. Toma valores '<','>','>=','<=','==' y '!=' y sus equivalentes verbales ('le','ge','leq','geq','eq' y 'neq').
- **Secured_enrollment_quota_id_value:** Valor a comparar según el criterio anterior
- **Applicant_characteristic_i_criteria:** Criterio a cumplir para aplicar el reordenamiento. Toma valores '<','>','>=','<=','==' y '!=' y sus equivalentes verbales ('le','ge','leq','geq','eq' y 'neq').
- **Applicant_characteristic_i_value:** Valor a comparar según el criterio anterior. Se espera que i tome valores según la cantidad de características de los postulantes.
- **Order_qj:** Se espera que j tome valores según la cantidad de cuotas de la asignación.

Siblings: Archivo indicando las relaciones familiares entre postulantes. Es necesario solo en el caso de indicar 'sibling_priority_activation'=True como kwarg. Esta base debe incluir las siguientes columnas.

- **Applicant_id:** Corresponde a un id de postulante.
- **Sibling_id:** Corresponde a un id de postulante.

Links: Archivo indicando las postulaciones en bloque. Es necesario solo en el caso de indicar 'linked_postulation_activation'=True como kwarg. Esta base debe incluir las siguientes columnas:

- **Applicant_id:** Corresponde a un id de postulante.
- **Linked_id:** corresponde a un id de postulante.

3. Proceso de transformación de variables

A continuación se describe el proceso de transformación de datos para obtener las tablas requeridas por el algoritmo (target), a partir de las tablas recibidas desde Perú (source).

Vacancies		
Target column	Transformation process	Source columns utilized
Program_id	Se crea una nueva tabla que relaciona cada uno de los programas ofrecidos en la tabla vacancies.csv con un identificador autoincremental.	Vacancies table: <ul style="list-style-type: none"> • Índice autoincremental
Quota_id	En las tablas actuales no se recibe información sobre cuotas, por defecto todos los programas ofrecen cuota 1	
Institution_id		Vacancies table: <ul style="list-style-type: none"> • localId
Grade_id		Vacancies table: <ul style="list-style-type: none"> • gradeId
Regular_vacancies		Vacancies table: <ul style="list-style-type: none"> • totalVacancyNna
special_1_vacancies		Vacancies table: <ul style="list-style-type: none"> • totalVacancyNnaNee

Applicants		
Target column	Transformation process	Source columns utilized
Applicant_id		Postulants table: <ul style="list-style-type: none"> • applicantId
Grade_id		Demand table: <ul style="list-style-type: none"> • gradeId

Special_assignment		Postulants table: <ul style="list-style-type: none"> priority
Secured_enrollment_program_id	No aplica.	
Secured_enrollment_quota_id	No aplica	
Applicant_characteristics_1	No aplica	

Applications		
Target column	Transformation process	Source columns utilized
Applicant_id		Demand table: <ul style="list-style-type: none"> postulantId
Program_id	Se relaciona el programa seleccionado con el programa ofrecido en la tabla vacancias.csv y se usa el identificador autoincremental creado.	Vacancies table: <ul style="list-style-type: none"> Índice autoincremental
Quota_id	No aplica, todos los estudiantes postulan por defecto a la quota 1	
Institution_id		Demand table: <ul style="list-style-type: none"> localId
Ranking_program		Demand table: <ul style="list-style-type: none"> order
Priority_profile_program	Perfil de prioridad al que pertenece según preferencia hermano. Para Perú existen solo dos perfiles de prioridad. <ul style="list-style-type: none"> Perfil 1: no tiene preferencia hermano Perfil 2: sí tiene preferencia hermano 	
Priority_number_quota	Según el perfil de prioridad, el número asignado a la quota. <ul style="list-style-type: none"> Perfil 1: priority number igual a 1 Perfil 2: priority number igual a 0 	

Priority_profiles		
Target column	Transformation process	Source columns utilized
Priority_profile	Identificador autoincremental	
Priority_qi	<ul style="list-style-type: none"> • Perfil 1: asigna un 1 • Perfil 2: asigna un 0 	
Priority_profile_sibling_transition	<ul style="list-style-type: none"> • Perfil 1: transiciona a perfil de prioridad 2 • Perfil 2: transiciona a perfil de prioridad 2 	

Siblings		
Target column	Transformation process	Source columns utilized
Applicant_id		Postulations table: <ul style="list-style-type: none"> • postulantId
Sibling_id	Se consideran hermanos a todos los estudiantes que comparten un mismo representante legal.	Postulations table: <ul style="list-style-type: none"> • postulantId

Links		
Target column	Transformation process	Source columns utilized
Applicant_id		Postulations table: <ul style="list-style-type: none"> • postulantId
Linked_id	Se incluyen todos los estudiantes que comparten un mismo representante legal y que además hicieron una postulación en grupo. Es decir, un valor "G" en el campo typeId para la tabla de postulaciones	Postulations table: <ul style="list-style-type: none"> • postulantId • typeId

La tabla quota_order no se incluye en el procesamiento de datos. Debido a la naturaleza del proceso de aplicación en Perú, todos los estudiantes aplican a una única quota común, por lo que el término quota no tiene utilidad y la tabla quota_order no contiene registros.

La tabla `priority_profiles` no se relaciona con los datos de aplicantes o vacantes. Esta tabla refleja el tratamiento de las políticas de asignación, como lo es el caso de la preferencia hermano. Por lo tanto, esta tabla se crea automáticamente durante la ejecución del algoritmo y no debe proveerse como input al algoritmo.