

# Diagnóstico Taller Data Science 2017-2

Profesores: Diego Caro & Eduardo Graells Universidad del Desarrollo

Nombre:

## Instrucciones

Lea atentamente cada uno de los 3 problemas, y programe la solución utilizando su lenguaje de programación favorito. Tiempo total: **40 minutos**.

Por favor indique el lenguaje de programación que usará: .

## Pregunta 1 (2 pt)

---

El Centro Cultural Gabriela Mistral tiene un programa de descuentos para jóvenes. El problema es que en la base de datos del Centro todos los socios son jóvenes! Al parecer, una actualización de Microsoft Windows alteró los registros, y cambió las fecha de nacimiento de cada socio por valores aleatorios.

El área de Data Science le pide programar una función `fechaValida(dia, mes, anio)` que devuelva  si el argumento es una fecha real, o  si no. Ejemplo, `fechaValida(32, 1, 2017)` devuelve  porque el "32 de enero" no es una fecha es válida, o `fechaValida(-18, 9, 2017)` que también devuelve  porque "-18 de Septiembre" tampoco lo es (no considere bisiestos). La fecha se entregará como tres números enteros para el día, mes y año.

Para este problema usted debe:

1. (0.3 pt) Indicar al menos 2 situaciones en las cuáles una fecha es inválida. Por ejemplo, se sabe que no existe el 32 de Enero.
2. (1 pt) Programar la función `fechaValida(dia, mes, anio)`.
3. (0.2 pt) Explique cómo su código determina la validez de una fecha.

Recuerde que:

- Febrero tiene 28 días
- Abril, Junio, Septiembre y Noviembre tiene 30 días
- Enero, Marzo, Mayo, Julio, Agosto, Octubre y Diciembre tienen 31 días

## Pregunta 2 (2 pts)

---

El profesor Rossa tiene problemas para usar el computador, así que normalmente calcula el promedio de notas de de sus alumnos usando papel y lápiz. Cuando el profesor entregó los promedios, Guru-guru se dió cuenta que su promedio no correspondía a sus calificaciones. La coordinadora académica de la Facultad le pidió a usted diseñar un programa que permita calcular el promedio de notas, para ayudar al profesor Rossa.

Si usted desea ayudar al profesor Rossa, y hacer justicia con Guru-Guru, resuelva lo siguiente:

1. (1.6 pt) Escriba un programa que calcule el promedio. Asuma que se le entrega un arreglo/lista/vector con  $n$  números, cada uno de ellos representando una nota, y que **todas las calificaciones tienen la misma ponderación**.
2. (0.4 pt) Explicar cómo funciona su programa.

**Nota:** Recuerde que el promedio está definido como  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , donde  $n$  es el número de evaluaciones del curso y  $x_i$  representa una nota particular.

## Pregunta 3 (3 pts)

---

Usted ha sido contratada por la Organización de las Naciones Unidas como consultora experta Data Science. Su primer trabajo es estudiar las economías de los países de acuerdo al nivel de desigualdad de ingresos. La desigualdad en los ingresos se puede cuantificar a través del Coeficiente de Gini. Este coeficiente varía entre 0 y 1, donde 0 indica igualdad absoluta y 1 desigualdad absoluta.

Su primera tarea como Data Scientist es reportar los 5 países con más desigualdad en los ingresos. Con este fin, usted debe hacer un programa que lea el archivo `gini_by_country.csv` y que imprima en pantalla estos 5 países.

**Nota:** el archivo `gini_by_country.csv` contiene varias líneas, una por país. Cada línea contiene dos campos separados por coma, donde el primer campo indica el país y el segundo campo indica su Coeficiente de Gini. Por ejemplo, la línea `Chile,0.454` indica que Chile tiene un Coeficiente Gini de 0.454 (fuente: <http://stats.oecd.org/Index.aspx?DataSetCode=IDD>)