

## Algoritmos y Estructuras de Datos

### Trabajo Práctico 2 (2024)

#### Implementación de un Sistema Simple de Gestión de Envíos por Correo (y Códigos Postales) - Versión 2.0

- Introducción:

El *Correo Central Argentino*, conociendo el éxito de la versión 1.0 desarrollada para el TP1, solicita ahora una versión 2.0 del programa que permita obtener estadísticas a partir de todas las operaciones y envíos del último mes. Al igual que en la versión anterior, los CP de Argentina y de los países vecinos de Argentina son como los del siguiente modelo ("L" significa *Letra*, y "N" significa *Número o Dígito (valores del 0 al 9)*):

Tabla 1:

País	Formato del CP	Observación
Argentina	LNNNNLLL	8 caracteres. La primera letra identifica a la provincia, los cuatro dígitos a la ciudad, y las tres letras finales al frente de manzana o paraje o casilla de correo.
Bolivia	NNNN	4 caracteres, todos dígitos. El número completo identifica a la ciudad o localidad.
Brasil	NNNNN-NNN	9 caracteres incluido el guion (el resto, todos dígitos). El primer dígito identifica una región, el segundo una subregión, del tercero al quinto un sector/subsector, y los tres últimos un sufijo de distribución.
Chile	NNNNNNN	7 caracteres, todos dígitos. Los tres primeros dígitos identifican a una comuna, y los cuatro últimos al frente de manzana.
Paraguay	NNNNNN	6 caracteres, todos dígitos. Los dos primeros identifican un departamento, el tercero y el cuarto a un distrito, y los dos últimos a un barrio o localidad.
Uruguay	NNNNN	5 caracteres, todos dígitos. Los dos primeros identifican un departamento (si el primero es 1, es la ciudad de Montevideo). Los tres últimos identifican una zona postal.

Igual que en la versión 1.0, solo a los efectos de este trabajo (*y sin que esto sea necesariamente real*), se asume que el Correo Central Argentino tiene aranceles por tipo de envío, por peso, y por destino de acuerdo a los siguientes cuadros (en los que solo consideraremos los tipos de envío que están enumerados, para simplificar el planteo de este TP):

a. Envíos nacionales (dentro de la propia Argentina):

Tabla 2:

Tipo de Envío	Id	Peso (p) en gramos	Precio (en pesos)
Carta Simple	0	$p < 20$	1100
Carta Simple	1	$20 \leq p < 150$	1800
Carta Simple	2	$150 \leq p < 500$	2450
Carta Certificada	3	$p < 150$	8300
Carta Certificada	4	$150 \leq p < 500$	10900
Carta Expresa	5	$p < 150$	14300
Carta Expresa	6	$150 \leq p < 500$	17900

**b. Envíos internacionales (fuera de la propia Argentina):**

Se toman los mismos tipos de envío que se muestran en la tabla anterior, pero se ajustan los precios de acuerdo a la tabla que sigue:

**Tabla 3:**

País destino	Precio (en pesos)
Bolivia, Paraguay, Uruguay (Montevideo)	+20%
Chile, Uruguay (no Montevideo)	+25%
Brasil (regiones 8 y 9)	+20%
Brasil (regiones 0, 1, 2 y 3)	+25%
Brasil (regiones 4, 5, 6 y 7)	+30%
Otros países	+50%

Solo a los efectos de simplificar el modelo, en este TP se entenderá que el envío se realiza hacia "Otro" país cuando el CP informado sea una cadena que no corresponde a ninguno de los países considerados en este TP. *No se preocupe de casos en los que algún país no contemplado pudiera tener un CP con el mismo formato de los países tomados como referencia aquí.*

Para esta versión 2.0 el programa deberá procesar los datos de muchos envíos (y no solo de uno, como en la version 1.0). Los datos de todos esos envíos no serán cargados por teclado, sino que vendrán almacenados en un *archivo de texto* llamado *envios.txt* que será provisto para su procesamiento.

Insistimos: el programa que tienen que realizar **no debe tomar datos desde el teclado por ninguna razón, ni debe tener un menú de opciones** (eso implicaría cargar por teclado la opción elegida). Tampoco debe haber en ninguna parte **nada** que solicite al usuario presionar una tecla para continuar. Todos los datos que el programa debe procesar (y solo esos datos) estarán en el archivo de texto *envios.txt* cuyo formato y contenido se describe más abajo.

Es MUY IMPORTANTE que ese archivo se llame *envios.txt* y que esté almacenado en la misma carpeta del proyecto que contiene al programa fuente que desarrollen. Si están haciendo pruebas con otros archivos, no hay problemas, pero ASEGÚRENSE de que cuando entreguen el programa para su calificación, la función *open()* esté abriendo exclusivamente un archivo llamado *envios.txt* Y NO OTRO.

El archivo de texto con los datos de entrada consistirá de una primera línea (que aparecerá por única vez) en la que de alguna forma no estándar se indicará la fecha y la hora de creación de ese archivo (una línea de este tipo en un documento, conteniendo fecha y hora de algún evento, se suele designar en el mundo informático como una *línea de marca temporal* (o *timestamp*)).

Y en alguna parte de esa línea, además, aparecerán dos caracteres de la forma "HC" (para indicar estructura de control rígida (*Hard Control*)) o de la forma "SC" (para indicar estructura de control ligera (*Soft Control*)). El tipo de estructura de control hace referencia a la forma en que se deben verificar las direcciones de destino de cada envío:

- **Hard Control (HC):** en cada envío del archivo de entrada, se debe controlar que la dirección de destino tenga solo letras y dígitos, y que no haya dos mayúsculas seguidas, y que haya al menos una palabra compuesta sólo por dígitos. Será considerado válido el envío solo si pasa la verificación indicada aquí.
- **Soft Control (SC):** cualquier envío es válido, sin importar el formato de la dirección de destino.

Algunos ejemplos (pero no los únicos posibles) de líneas de *timestamp* para este trabajo son los siguientes:

- ✓ 21 de mayo de 2023 – 14 hs 52 min – HC
- ✓ SC – 18:15 – 10/06/2020
- ✓ 07:10 – HC – 2024-07-08

En definitiva: una línea timestamp contendrá datos de fecha, hora y tipo de estructura de control, pero sin seguir un formato fijo ni estándar. Los dos grupos de caracteres "HC" y "SC" siempre aparecerán una y solo una vez, y solo uno de ellos, pero en cualquier lugar dentro de esa línea.

El resto de las líneas del archivo de entrada contendrá los datos de un envío, todos juntos conformando una cadena de 31 caracteres de largo (distribuidos como se explica más abajo), siempre con el mismo formato. Cada línea se separa de la que sigue con salto de línea.

Si se tienen registrados los datos de 20 envíos, entonces el archivo tendrá la línea de timestamp y luego otras 20 líneas, una por cada envío. En cada línea aparecerán, estrictamente en este orden, los siguientes datos (que son los mismos que en la versión 1.0 se cargaban por teclado):

- ✓ **Caracteres del 0 al 8 (o sea, 9 caracteres):** *El código postal.* Si envío es hacia cualquier país que no sea Brasil, entonces los primeros caracteres que sean necesarios serán espacios en blanco. De esa forma, en el archivo el código postal estará siempre conformado por 9(nueve) caracteres y los blancos al inicio eventualmente deberán ser removidos antes de procesar el envío. Igual en la versión 1.0, cualquier combinación de nueve caracteres es válida, pero no necesariamente todas corresponden a uno de los países en los que se centra este trabajo.
- ✓ **Caracteres del 9 al 28 (o sea, 20 caracteres):** *La dirección física del destino.* Una cadena de caracteres indicando la dirección concreta de entrega. Esta cadena es la que debe ser controlada si en la línea de timestamp aparece la secuencia "HC". La cadena puede tener espacios en blanco al final para completar los 20 caracteres de largo, y esos blancos extra deben ser removidos antes de procesar la cadena.
- ✓ **Caracter 29:** *El tipo de envío.* Un número entero entre 0 y 6 que indica alguno de los siete tipos posibles (ver columna "Id" en la tabla de tipos de envío - Tabla 2).
- ✓ **Caracter 30:** *La forma de pago.* Un número entero que indica alguno de los dos siguientes tipos de pago: (1: efectivo, 2: tarjeta de crédito).

Ejemplo: En el siguiente modelo mostramos qué forma podría tener un pequeño archivo de texto como el indicado:

07:10 - HC - 2024-07-08	
97223-374Sao Gabriel 568	11
D2345LHGSan Martin 856	02
X5389GKSIndependencia 9863	42
8475SAntonio Sucre 83	51
13774-244Sao Corrado 5587	31
764114Paraguari 1441	42
AD9541Libertad 5277	61

En el modelo anterior la primera línea de todas es la de *timestamp*, y los caracteres de tipo de control son "HC" (aproximadamente al centro de la línea).

Luego siguen otras 7 líneas (una por envío registrado):

1. La primera de esas siete es de un envío a Brasil, tipo envío 1, forma de pago 1. Note en esta línea (y en todas las que siguen) los espacios en blanco al final de la dirección física y antes del dígito de tipo de envío, para completar los 20 caracteres asignados a la dirección física.
2. La segunda es de un envío a Argentina, tipo 0, forma de pago 2. Note el espacio en blanco delante del código postal.
3. La tercera es de un envío a Argentina, tipo 4, forma de pago 2. Note el espacio en blanco delante del código postal.
4. La cuarta es de un envío a Bolivia, tipo 5, forma de pago 1. Note ahora los cinco espacios en blanco delante del código postal.

5. La quinta es de un envío a Brasil, tipo 3, forma de pago 1.
6. La sexta es de un envío a Paraguay, tipo 4, forma de pago 2. Note ahora los tres espacios en blanco delante del código postal.
7. La séptima es de un envío a Otro país, tipo 6, forma de pago 1. Note que el código postal tiene 6 caracteres, pero NO ES de Paraguay ya que contiene letras. Y otra vez, note los tres espacios en blanco delante del código postal.

#### ● **Requerimiento:**

Se pide desarrollar un programa que en base a todo lo anterior, procese los datos de un archivo con el formato indicado, usando un ciclo para levantar una por una las líneas del archivo, y muestre finalmente los siguientes resultados (y **estrictamente** estos resultados):

- **Resultado 1 (r1):** *Tipo de control de direcciones detectado en la línea de timestamp:* Mostrar la cadena “Hard Control” o la cadena “Soft Control” de acuerdo a si en la línea de timestamp del archivo se detectó la secuencia “HC” o la secuencia “SC” respectivamente. Se aceptarán como válidas **solamente** las cadenas “Hard Control” o “Soft Control”, en cualquier combinación de minúsculas o mayúsculas.
- **Resultados 2 y 3 (r2 y r3):** Si el tipo de control de direcciones era HC, mostrar *la cantidad de envíos con dirección válida* (r2), y *la cantidad de envíos con dirección no válida* (r3). Recordar que si el tipo de control es HC, entonces se considera que una dirección es válida si cumple que tiene solo letras y dígitos, y que no tenga dos mayúsculas seguidas, y que tenga al menos una palabra compuesta sólo por dígitos.
- Si el tipo de control de direcciones era SC, mostrar entonces *la cantidad total de envíos* (r2) (*que en este caso significará que todas las direcciones son válidas*), y *muestre un cero en el resultado r3* (*indicando con ello que no hay direcciones inválidas. PERO MUESTRE EL CERO*).
- **Resultado 4 (r4):** *El importe final acumulado por todos los envíos.* Si el tipo de control era HC, hay que acumular solo los importes finales de los envíos con direcciones válidas. Si el tipo de control era SC, hay que acumular todos los importes finales, de todos los envíos.
- **Resultados 5, 6 y 7 (r5, r6 y r7):** La cantidad de envíos como Carta Simple (r5), como Carta Certificada (r6) o como Carta Expresa (r7) (tres contadores, uno para cada tipo de carta). De nuevo, si el tipo de control era HC, hay que contar solo los envíos con direcciones válidas. Si el tipo de control era SC, hay que contar todos los envíos.
- **Resultado 8 (r8):** Tomando los tres conteos del punto anterior, indicar *cuál es el tipo de carta con mayor cantidad de envíos*. Las cadenas a mostrar deben ser de la forma “Carta Simple” o “Carta Certificada” o “Carta Expresa”, y solo alguna de esas tres, en cualquier combinación de minúsculas y mayúsculas. Si el mayor conteo estuviese repetido, informe el tipo de carta más sencillo entre los que se repiten (ejemplo: si el mayor conteo fuese el de las simples y fuese igual al de las expresas, informe que el tipo mayor es “Carta Simple”).
- **Resultados 9 y 10 (r9 y r10):** Indicar *cuál es el CP del primero de los envíos del archivo* (r9), y *la cantidad de veces que apareció ese CP en el archivo* (r10). Para los resultados r9 y r10, no considerar el tipo de control de direcciones (dar los resultados sin importar si el tipo de control era HC o SC).
- **Resultados 11 y 12 (r11 y r12):** El *menor importe final pagado por un envío a Brasil* (r11) y *el CP de ese envío* (r12). Si ese menor importe final se repite en más de un envío, tomar el primero que haya encontrado para el menor importe e informar el CP de ese. Para los resultados r11 y r12, no considerar el tipo de control de direcciones (dar los resultados sin importar si el tipo de control era HC o SC).
- **Resultado 13 (r13):** El *porcentaje que representa la cantidad total de envíos al exterior sobre la cantidad total de envíos*. Si el tipo de control era HC, hay que contar solo los envíos con direcciones válidas. Si el tipo de control era SC, hay que contar todos los envíos. El porcentaje debe mostrarse en

forma truncada, sin decimales (realizar primero la multiplicación, y luego la división. Haga el cálculo con división real, y al final trunque el resultado usando la función `int()`). Prevea que el programa no se clave por una eventual división por cero: si el porcentaje no puede calcularse, muestre un cero como resultado.

- **Resultado 14 (r14):** El *monto final promedio pagado por los envíos a la provincia de Buenos Aires*. Si el tipo de control era HC, hay que sumar solo los precios finales de los envíos con direcciones válidas. Si el tipo de control era SC, hay que sumar los precios finales de todos los envíos. El promedio debe mostrarse en forma truncada, sin decimales (haga el cálculo con división real, y al final trunque el resultado usando la función `int()`). Prevea que el programa no se clave por una eventual división por cero: si el promedio no puede calcularse, muestre un cero como resultado.

#### ● Consideraciones especiales:

El desarrollo de este TP2 debe realizarse exclusivamente en base a temas y prácticas presentados en las Fichas 1 a 13. El programa debe abrir el archivo de texto que se provea, procesar ese archivo línea por línea mediante un ciclo, usar funciones parametrizadas y con retorno de valores en toda situación posible, y disponer de una función de entrada correctamente planteada. El uso de funciones en estas condiciones es obligatorio: el planteo del programa en base a un script gigante, sin funciones de ningún tipo o con funciones totalmente elementales, será motivo de aplazo inmediato.

En la Ficha 11 y en las clases prácticas de materia, se irán presentando temas adicionales necesarios para el desarrollo completo de este trabajo (incluyendo la forma de procesar línea por línea un archivo de texto). Mientras tanto, los estudiantes pueden ir investigando por sus propios medios la forma de hacer lo que necesiten hacer. Y por supuesto, siempre pueden preguntar en clases o por medio del foro para consultas sobre el TP2.

Junto con este enunciado, se provee un archivo de prueba llamado [envios25.txt](#) con la línea de timestamp y luego 25 líneas con registros de prueba, para que los estudiantes puedan testear sus programas y validar básicamente lo que están haciendo. Para ese archivo en particular, los resultados que deberían obtener son los siguientes:

```
(r1) - Tipo de control de direcciones: Hard Control
(r2) - Cantidad de envios con direccion valida: 17
(r3) - Cantidad de envios con direccion no valida: 8
(r4) - Total acumulado de importes finales: 152627
(r5) - Cantidad de cartas simples: 9
(r6) - Cantidad de cartas certificadas: 3
(r7) - Cantidad de cartas expresas: 5
(r8) - Tipo de carta con mayor cantidad de envios: Carta Simple
(r9) - Codigo postal del primer envio del archivo: 03X3076
(r10) - Cantidad de veces que entro ese primero: 2
(r11) - Importe menor pagado por envios a Brasil: 2250
(r12) - Codigo postal del envio a Brasil con importe menor: 07439-412
(r13) - Porcentaje de envios al exterior sobre el total: 60
(r14) - Importe final promedio de los envios a Buenos Aires: 10900
```

Se proveen además otros dos archivos de texto llamados [envios100SC.txt](#) y [envios100HC.txt](#): se trata en realidad del mismo archivo, con las mismas 100 líneas de datos de prueba, pero solo variando la línea de timestamp: en el primero, la secuencia de control de direcciones es "SC", y en segundo es "HC". Pero fuera de ese pequeño detalle, los datos de ambos archivos son los mismos. Los resultados esperados para el archivo [envios100SC.txt](#) son los siguientes:

```
(r1) - Tipo de control de direcciones: Soft Control
(r2) - Cantidad de envios con direccion valida: 100
(r3) - Cantidad de envios con direccion no valida: 0
(r4) - Total acumulado de importes finales: 943197
(r5) - Cantidad de cartas simples: 42
(r6) - Cantidad de cartas certificadas: 34
(r7) - Cantidad de cartas expresas: 24
(r8) - Tipo de carta con mayor cantidad de envios: Carta Simple
(r9) - Codigo postal del primer envio del archivo: 247X
(r10) - Cantidad de veces que entro ese primero: 1
(r11) - Importe menor pagado por envios a Brasil: 1287
(r12) - Codigo postal del envio a Brasil con importe menor: 65090-152
(r13) - Porcentaje de envios al exterior sobre el total: 86
(r14) - Importe final promedio de los envios a Buenos Aires: 16110
```

Y los resultados esperados para el archivo [envios100HC.txt](#) son los siguientes:

```
(r1) - Tipo de control de direcciones: Hard Control
(r2) - Cantidad de envios con direccion valida: 63
(r3) - Cantidad de envios con direccion no valida: 37
(r4) - Total acumulado de importes finales: 571649
(r5) - Cantidad de cartas simples: 31
(r6) - Cantidad de cartas certificadas: 15
(r7) - Cantidad de cartas expresas: 17
(r8) - Tipo de carta con mayor cantidad de envios: Carta Simple
(r9) - Codigo postal del primer envio del archivo: 247X
(r10) - Cantidad de veces que entro ese primero: 1
(r11) - Importe menor pagado por envios a Brasil: 1287
(r12) - Codigo postal del envio a Brasil con importe menor: 65090-152
(r13) - Porcentaje de envios al exterior sobre el total: 54
(r14) - Importe final promedio de los envios a Buenos Aires: 16110
```

Repetimos para evitar sorpresas después: **El uso de funciones correctamente parametrizadas y con retorno de valores si fuese aplicable, es obligatorio: el planteo del programa en base a un script gigante, sin funciones de ningún tipo o con funciones totalmente elementales, será motivo de aplazo inmediato.**

Al igual que el TP1, este trabajo también será calificado en forma automática (pero se controlará el cumplimiento del uso de funciones y el esquema general exigido): Tanto las entradas (que vendrán desde un archivo de texto) como las salidas, deben atenerse estrictamente a lo esperado. El archivo de texto de entrada debe ser leído asumiendo el orden correcto de datos que contiene, y también es totalmente exigible el orden de las salidas, tal cual como se muestran en la captura para el archivo envios25.txt de la página anterior.

No muestren nada más **en ninguna otra parte del programa**: Los únicos print() que el programa debe tener son los que están indicados en el enunciado para generar la secuencia de salidas pedidas. Y **nada más** que esos print() que se ven en la secuencia más abajo en esta misma página. No pongan títulos. No pongan mensajes de bienvenida. Ni saludos. Ni agradecimientos. No cambien caracteres en los mensajes de salida que indicamos más abajo. No agreguen líneas en blanco. No agreguen ningún print() de control. Nada. Limítense **ESTRICTAMENTE** a los print() que se indican más abajo.

Mostramos ahora la secuencia de instrucciones *print()* que sus programas deberían tener al final (copien y peguen al final de su código fuente toda la secuencia de salidas que sigue), y respeten a rajatabla **el orden de esas salidas, el formato, los mensajes y la cantidad de líneas** que les estamos mostrando. En la secuencia de abajo, lo único que pueden cambiar son los nombres de las variables en los que se suponen almacenados los resultados (obviamente, con esos nombres cada programador decide):

```
print(' (r1) - Tipo de control de direcciones:', control)
print(' (r2) - Cantidad de envios con direccion valida:', cedvalid)
print(' (r3) - Cantidad de envios con direccion no valida:', cedinvalid)
print(' (r4) - Total acumulado de importes finales:', imp_acu_total)
print(' (r5) - Cantidad de cartas simples:', ccs)
print(' (r6) - Cantidad de cartas certificadas:', ccc)
print(' (r7) - Cantidad de cartas expresas:', cce)
print(' (r8) - Tipo de carta con mayor cantidad de envios:', tipo_mayor)
print(' (r9) -Codigo postal del primer envio del archivo:', primer_cp)
print(' (r10) - Cantidad de veces que entro ese primero:', cant_primer_cp)
print(' (r11) - Importe menor pagado por envios a Brasil:', menimp)
print(' (r12) -Codigo postal del envio a Brasil con importe menor:', mencp)
print(' (r13) - Porcentaje de envios al exterior sobre el total:', porc)
print(' (r14) - Importe final promedio de los envios Buenos Aires:', prom)
```