

## Tarea 1

Ricardo Esteban Schilling Broussaingaray

Tomás Alberto Soto Gonzales

Dunkan Nicolás Torres Gonzalez

Jessica Yanín Hormazabal Pardo

IIC2713 - 1 - Sistemas de información

Grupo 18

Contraseña Excel - "g18"

2020

## Tabla de Contenido

<b>1</b>	<b>Limpieza de Datos</b>	<b>1</b>
	<b>Limpieza de Datos</b>	<b>1</b>
1.1	Libro Conteo General por País . . . . .	1
1.1.1	Casos Confirmados . . . . .	1
1.1.2	Fallecidos / 100 Casos . . . . .	1
1.1.3	Recuperados / 100 Casos . . . . .	2
1.1.4	Fallecidos / 100 Recuperados . . . . .	2
1.2	Datos de Worldometer . . . . .	2
1.2.1	Mongolia y Mauritius . . . . .	2
1.2.2	Población . . . . .	3
1.2.3	Casos Totales . . . . .	3
1.2.4	Tests totales . . . . .	3
<b>2</b>	<b>Análisis Chile</b>	<b>5</b>
2.1	Diferencia entre conteo diario por país y Worldometer . . . . .	5
2.2	Preguntas Finales . . . . .	5
2.2.1	¿En qué mes se produjo el mayor incremento en los casos de contagio (respecto al mes pasado)? ¿Cuánto fue? . . . . .	5
2.2.2	¿Qué día ocurrió el <i>peak</i> de nuevos fallecidos? . . . . .	6

	II
2.2.3 ¿Qué mes ocurrió el día <i>peak</i> de fallecidos?¿Cuánto fue el peak? . . . . .	6
2.2.4 ¿Qué mes tuvo el mayor número de contagiados?¿Cuántos fueron? . . . . .	6
<b>3 AnalisisMundial</b>	<b>7</b>
3.1 Cambio porcentual entre 10 países . . . . .	7
3.1.1 ¿Qué país presenta mayor cambio porcentual?¿Y el menos? . . . . .	7
3.1.2 A partir de la información sobre la duración del virus, concluye cuánto tiempo aproximadamente hay desfase entre los valores comparados . . . . .	7
3.2 Analizar y responder . . . . .	7
3.2.1 ¿Qué continente posee el mayor números de contagios?¿Cuál el menor? Explica a qué se puede deber esta diferencia. . . . .	7
3.2.2 ¿En qué país se registra la mayor cantidad de recuperados? . . . . .	8
3.2.3 ¿En qué continente se registra la mayor cantidad de fallecidos? . . . . .	8
3.2.4Cuál es el país con mayor porcentaje de casos activos? Mencione de qué nos puede servir esta estadística al momento de analizar la realidad de la pandemia en dicho país . . . . .	8
<b>4 Seguridad</b>	<b>9</b>

## 1. Limpieza de Datos

### 1.1. Libro Conteo General por País

Antes de comenzar la limpieza, lo primero que se realizó fue una comprobación de los datos entregados. Por lo cual se hizo uso de los Formatos Condicionales para dejar de forma más evidente, visualmente hablando, de los posibles errores que se pueden encontrar en cada celda.

Se hace el supuesto de que existen columnas con datos en los cuales se puede confiar, estos serán utilizados para confirmar otras columnas y, en caso de que se encuentre una contradicción a este supuesto, se tomarán las medidas necesarias para resolverlo.

A continuación se dará una breve explicación de cómo se comprobaron ciertos datos y se presentarán las formulas utilizadas para la validación en cada uno de ellos:

#### 1.1.1. Casos Confirmados

Esta columna en específico puede comprobarse con una sumatoria entre las columnas 'Fallecidos totales', 'Recuperados totales' y 'Casos activos'.

La condición utilizada trabaja desde el intervalo  $D2 : D183$  (Toda la columna) con la condición de que el valor de celda sea igual a  $F2 + H2 + J2$

No se encontraron incongruencias en la columna, por lo que no se hacen modificaciones.

#### 1.1.2. Fallecidos / 100 Casos

A diferencia de la tabla anterior esta considera la columna de 'Fallecidos' (C) y 'Confirmados' (B), trabajando con la condición de que la celda sea igual a  $REDONDEAR((C * 100) / B)$ , de forma generalizada.

No se encontraron incongruencias en la columna, por lo que no se hacen modificaciones.

### 1.1.3. *Recuperados / 100 Casos*

Asemejándose a lo anterior, la formula se mantiene pero cambiamos la columna 'Fallecidos' por 'Recuperados' (D), dejándonos como condición que la celda sea, de forma generalizada, igual a  $REDONDEAR((D * 100) / B)$ .

No se encontraron incongruencias en la columna, por lo que no se hacen modificaciones.

### 1.1.4. *Fallecidos / 100 Recuperados*

**Esta columna fue eliminada de la base de datos.** Al existir países donde su conteo de recuperados eran iguales a cero, esto producía indeterminación en la división que se intentaba implementar.

Entonces, comparando los costos entre perder la información de países que me será irrecuperable en caso de necesitarla, versus una columna que puede re-calcularse en cualquier momento, preferimos eliminarla para que no traiga problemas al momento de realizar las operaciones en la parte de Macros.

## 1.2. Datos de Worldometer

En este libro nos encontramos con 3 casos principales del cual nos hicimos cargos más un error de *typo*. El primero tiene que ver con la información entregada en la columna de Población, otro con la cantidad de Casos totales y Tests totales realizados, además de una confusión que se hizo entre Mongolia y Mauritius.

### 1.2.1. *Mongolia y Mauritius*

En la hoja original entregada podemos notar que las líneas 160 y 161 presentan dos veces al país de Mauritius pero, en distintos continentes (África y Asia respectivamente). Por lo cual nos dimos el trabajo de revisar la base original entregada por *Worldometer* a través del link que se nos proporcionó en el enunciado del trabajo.

Al ir comparando la información del documento original y el resto de columnas de ambos 'Mauritius' pudimos dilucidar que el país al cual se referían en la línea 161 es Mongolia.

### 1.2.2. Población

Al comprobar que la columna población sea consistente con los otros datos. Sin embargo, La información proporcionada por la hoja no era la suficiente para hacer una relación directa entre los datos. Sin embargo, al calcular las relaciones 'Casos totales' y 'Fallecidos' entre 1M de población, pudimos percatarnos de que los valores no coincidían con la población entregada.

Por lo anterior, nos dirigimos a la base original de *Worldometer* y actualizamos cada una de las poblaciones por la correspondiente a esta base de datos, ya que todos los valores que poseemos concuerdan con la base original, sin contar la columna Población.

Al realizar el cambio, las relaciones entre columnas comienzan a tener sentido y se da por solucionado el problema.

### 1.2.3. Casos Totales

A pesar de la confianza que parece existir en la base de *Worldometer*, existen datos faltantes los cuales fueron omitidos ó considerados como cero, sin embargo, esta última aceveración pareciera ser falsa, ya que la cantidad de 'Casos totales' no parece concordar con la información entregada por los datos que sí poseemos, como lo son 'Fallecidos totales'.

Por dar un ejemplo, *Spain* cuenta con 354,530 'Casos activos', sin embargo, 'Fallecidos totales' son 28,500, 'Casos críticos' 617 y el resto de información es vacío, por lo tanto, si se considera como cero, la cantidad de casos activos es mayor a la información presentada.

Razón por la cual se decidió usar la información proporcionada por la hoja 'Conteo General por País' para 'rellenar' estos espacios en blanco, cuidando de que los datos ingresados sean una aproximación confiable de la presentada por la hoja actual.

### 1.2.4. Tests totales

**Se eliminaron cerca de 22 países de la base** debido a que no presentaban 'Tests realizados', lo cual perjudicaba en la realización de las Macros en la parte 3 del trabajo.

Sin considerar que la información de la columna era necesaria para un paso posterior al de la limpieza, podríamos considerar que no estamos perdiendo información valiosa ya que la hoja Conteo

General por país es una buena aproximación de la información presentada en la presente hoja.

Por lo cual simplemente decidimos eliminar lo que no nos permitía realizar el trabajo.

## 2. AnalisisChile

### 2.1. Diferencia entre conteo diario por país y Wordometer

Empezando por el hecho de que realmente no conocemos la procedencia de los datos entregados por el archivo *conteo diario por país*, luego de un poco de investigación, *Wordometer* es una asociación completamente independiente, sin fines de lucro, que proporciona estimaciones y estadísticas en tiempo real.

En su web dice que para los datos del COVID-19, se recopilaron los datos desde informes oficiales, sacados directamente de los canales de comunicación del gobierno o, incluso, de forma indirecta, de fuentes que consideraron confiables. Sus analistas manejan información de más de 5,000 fuentes, por lo que no es alocado decir que el **cuando** fue extraída la información que se nos proporcionó es más importante de lo que parece.

Suponiendo que el origen del conteo diario por país es información proporcionada específicamente por cada país, tiene sentido que en la hoja *Wordometer* falte información o, mejor dicho, no sea tan precisa en comparación a la proporcionada por los gobiernos.

Sin embargo, no nos malentiendan, esta diferencia no desacredita la información proporcionada por *Wordometer*, de hecho, según lo investigado, los mismos gobiernos hacen uso de esta aplicación para mantenerse al tanto del avance del virus alrededor del mundo.

Pero, como se mencionó anteriormente, al ser una actualización constante a **tiempo real**, una recopilación de datos tradicional no podrá compararse nunca con la precisión que intenta llegar la web de *Wordometer*.

### 2.2. Preguntas Finales

**2.2.1. ¿En qué mes se produjo el mayor incremento en los casos de contagio (respecto al mes pasado)? ¿Cuánto fue?**

El mes que produjo el mayor incremento de nuevos casos respecto al mes pasado es mes de **Mayo** con una diferencia de 90990 nuevos casos respecto al mes anterior.



### ***2.2.2. ¿Qué día ocurrió el peak de nuevos fallecidos?***

El 17/07/2020 es el día en que ocurrió el *peak* de nuevos fallecidos en Chile. con un total de 1,057 fallecidos.

### ***2.2.3. ¿Qué mes ocurrió el día peak de fallecidos? ¿Cuánto fue el peak?***

En el mes de Julio, con 1,057 fallecidos...

### ***2.2.4. ¿Qué mes tuvo el mayor número de contagiados? ¿Cuántos fueron?***

El mes de Junio, específicamente el 06/06/2020 aparecieron 13,990 nuevos casos. Superando el mes anterior por 872 casos.

### 3. AnalisisMundial

#### 3.1. Cambio porcentual entre 10 países

##### 3.1.1. *¿Qué país presenta mayor cambio porcentual? ¿Y el menos?*

El país que presenta mayor cambio porcentual es Egipto, con 40,36\$ y el menor lo tiene Irlanda con 0%

##### 3.1.2. *A partir de la información sobre la duración del virus, concluye cuánto tiempo aproximadamente hay desfase entre los valores comparados*

Considerando que Woeldometer es una aplicación que revela la información a tiempo real, el que exista un país con un 0% de cambio porcentual da la ilusión de que el tiempo de desfase es nulo.

Es idílico pensar que la información obtenida por el Conteo general por país se de a la velocidad con la que Worldometer accede a ella. Un supuesto razonable es que, debido a la emergencia a nivel mundial, no exista un intervalo menor de 24 ni mayor a 48-50 horas de desfase entre los valores comparados.

#### 3.2. Analizar y responder

##### 3.2.1. *¿Qué continente posee el mayor números de contagios? ¿Cuál el menor? Explica a qué se puede deber esta diferencia.*

El continente con el mayor número de contagios es North América con 5913739 contagios, el menor es Austria/Oceanía con no más de 21649.

Esta diferencia en primer lugar se puede deber a que solo tenemos registrados 4 países de Austria/Oceania, versus los 23 de North América, de los cuales tenemos a Estados Unidos (USA), quien es el país con mayor número de contagios, entre todos los países.

Podríamos decir que la diferencia de números se debe a la cantidad de países involucrados en el conteo pero, nos hace investigar sobre qué medidas a tomado el continente con el menor número

de contagios para combatir los casos de COVID-19.

### ***3.2.2. ¿En qué país se registra la mayor cantidad de recuperados?***

Sorprendentemente, el país que registra mayor cantidad de recuperados es Estados Unidos, con 2,576,668 casos. Sin embargo, si vemos el porcentaje respecto a los casos totales no es más que el 50 % de los casos totales. Al menos dos países tienen el 100 % de casos recuperados, los cuales son Dominica y Greenland, aunque también son países con no más de 18 y 14 casos, respectivamente.

### ***3.2.3. ¿En qué continente se registra la mayor cantidad de fallecidos?***

*North America* es el continente que se registra con la mayor cantidad de fallecidos.

### ***3.2.4. ¿Cuál es el país con mayor porcentaje de casos activos? Mencione de qué nos puede servir esta estadística al momento de analizar la realidad de la pandemia en dicho país***

Sweden es el país con mayor porcentaje de casos activos, con un valor del 92,92 %.

Esta información es útil ya que nos alerta acerca de la situación del país, de hecho, si le echamos un vistazo a la cantidad de recuperados nos damos cuenta de que es de 0. Es decir, que durante todo este tiempo la gente infectada a muerto por el virus o se mantiene con la enfermedad. Esto pone al país en una situación crítica y de emergencia respecto a otros países cuyos porcentajes son menores al 20 %, por decir un caso.

Los números nos están avisando de que tenemos que prestarle atención a las medidas que está ejecutando Sweden, puede que necesite recursos humanos o médicos, quizás no están cumpliendo la cuarentena, todo esto nos abre paso a un "Tenemos que investigar/Poner el foco en países como Sweden, para evitar que se siga esparciendo la enfermedad alrededor del mundo

## 4. Seguridad

### **Contraseña: g18**

Las medidas que tiene Excel para proteger los datos se basan mayoritariamente en el uso de una contraseña para encriptar la información contenida en el archivo, permitiendo que solo alguien con la contraseña pueda ver los contenidos.

Otro tipo de protección consiste similarmente en la aplicación de una contraseña para modificar un archivo específico, a modo de proteger la integridad de los datos que esta pueda contener. Algo muy similar que se puede hacer es marcar un archivo como "final", evitando que cualquier cambio se pueda realizar.

Finalmente, similar a las otras opciones se puede usar una contraseña para evitar cambios en la estructura del archivo, ya sea evitando la agregación de nuevas páginas, que estas se escondan o de manera mas específica restringiendo las acciones que puede realizar un usuario dentro del archivo.

Como se puede concluir, todos los tipos de restricciones que provee Excel dependen principalmente de una contraseña.