

ICO 187 ANÁLISIS DE DATOS

GUÍA TAREA GRUPO N01

Año: 2021

Profesor: Sebastián Egaña

1. Instrucciones generales

1. El informe debe ser entregado en formato word, considerando también el envío de las planillas relacionadas con su trabajo. Considere la incorporación de la data bruta descargada (sin modificaciones) como también la data trabajada por usted. Todo debe venir contenido en un archivo zip.
2. El archivo debe ser enviado al correo del profesor: sebastianeganasa@santotomas.cl, hasta las 23:59 del día de entrega. La fecha de entrega es el **04 de mayo del 2021**.
3. Los grupos deben ser de 3 a 4 personas. Sin excepciones.
4. Se debe enviar un correo al profesor, especificando los integrantes y adjuntando el set de datos a trabajar hasta el domingo 25 de abril de 2021. En caso de tener personas en el grupo que no trabajen o no cooperen, se le debe informar al profesor.
5. Solo deben utilizarse funciones para realizar sus cálculos; no se acepta la utilización de valores “pegados”, tablas dinámicas, utilización de análisis de datos o agrupamientos raros en Excel. Considere que en caso de reemplazos, esto no es necesario pero se deben especificar dichos procedimientos.

2. Puntos a desarrollar

1. Se debe obtener un set de datos de las siguientes páginas:

- [Observatorio Logístico](#)
- [Conjunto de datos del Gobierno](#)

Dicho set de datos, debe ser enviado al docente para su aprobación (como se indica en el punto 3 de las instrucciones generales). Por ejemplo, se espera que utilicen set de datos con una cantidad considerable de columnas y filas, y que tengan un cierto grado de complejidad para el análisis. Por encima de esto, los datos deben tener una orientación pública.

Considere la posibilidad de obtener datos de otro tipo de fuente, en caso de cumplir con los requisitos anteriores.

(Sin puntos, pero paso necesario)

2. Especifique si tuvo o no que desarrollar algún procedimiento de limpieza, ordenamiento o transformación de los datos. Sea detallista en esto. En el caso de no realizar procedimiento alguno, fundamente.

(10 puntos)

3. Detalle en un cuadro, el nombre de cada variable y el tipo de dato de dicha variable.

(10 puntos)

4. Presente en un análisis del set de datos, que pueda dar una imagen inicial sobre los datos. Considere por ejemplo desarrollar estadísticas descriptivas generales, como también agrupadas.

El análisis no refiere a solo la incorporación de cifras. Debe intentar desarrollar un texto que explique o hable de ciertas características de los datos dentro del informe. Se premiará el desarrollo de visualizaciones según lo que se ha conversado en el curso sobre el tema (buenas visualizaciones).

Intente generar algunas preguntas y respuestas sobre el set de datos.

(30 puntos)

5. Considere la utilización de otro set de datos que podría contribuir al análisis del set de datos elegido. Dicho set de datos no deberá ser aprobado por el docente.

Se espera que este set de datos contribuya al punto anterior.

(20 puntos)

6. Desarrolle un apartado en donde haga el análisis sobre la existencia o no de datos anómalos. Por ejemplo, si tengo edades de personas, no deberían existir valores negativos, ni cero, ni mayores a 120. Su explicación debe desarrollar el porque cree que existen o no dichos valores en la data.

(10 puntos)

- Aparte: Se valorará de manera extra el realizar algún tipo de conexión compleja en los datos, como también el agregar valor al informe.