

**Juan Carlos** [25 oct. 2021 a la(s) 13:31](https://epirhandbook.slack.com/archives/C02794V09SB/p1635186676001900)

**HILO #1 TERMINOLOGÍA**En este hilo podemos añadir las dudas sobre traducción de términos. Lo empiezo con uno reciente e iremos añadiendo las dudas en la traducción:Tidy data. En el primer borrador está traducido como datos ordenados. [@Ximena](https://epirhandbook.slack.com/team/U025NCT1XSB) sugiere el termino ‘prolijo’ creo que se ajusta mejor. ‘Tidy’ significa mas que ‘ordenado’.  
Aunque creo que tienes razón, en el glosario de R 4 data science lo traduce como datos ordenados. He visto que en la [wikipedia](https://es.wikipedia.org/wiki/Tidy_data" \t "_blank) también y he visto otros documentos con la misma traducción. Así que si te parece, salvo que creas que el uso de prolijo está más extendido, lo podemos dejar en "Datos ordenados", refiriéndose ss  Conjuntos de datos 'tidy' están ordenados de tal manera que cada variable es una columna y cada observación (o caso) es una fila. Con algunas variantes, he encontrado esta frase en los otros documentos. (editado)

[Wikipedia](https://es.wikipedia.org/)[**Wikipedia**](https://es.wikipedia.org/)

[**Tidy data**](https://es.wikipedia.org/wiki/Tidy_data)

Los datos ordenados o 'tidy data'  son los que se obtienen a partir de un proceso llamado 'data tidying' u ordenamiento de datos.[1]​  Es uno  de los procesos de limpieza importantes durante procesamiento de grandes datos o 'big data' y es un paso reconocido en la ciencia de datos.  Los conjuntos de datos ordenados tienen una estructura que facilita el trabajo;  son sencillos de manipular, modelar y visualizar.  Conjuntos de datos 'tidy' están ordenados de tal manera que cada variable es una columna  y cada observación (o caso) es una fila.[1]​[2]  
Los datos ordenados proporcionan estándares y conceptos para los datos que limpian, y con el dato ordenado allí es ninguna necesidad de empezar de arañazo… Mostrar más

8 respuestas



**Ximena**  [hace 12 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1635245800012900?thread_ts=1635186676.001900&cid=C02794V09SB)

Gracias por compartir esto Juan Carlos, si los expertos en ciencia de datos hispanoparlantes usan 'datos ordenados' pues adelante, usamos eso. Yo no leo mucho de estas cosas en Español asique aprendo :cara_ligeramente_sonriente: Gracias a este trabajo creo que mejoro mi Español ya que no lo uso nunca...desde hace muchos años. Disculpen si ha veces escribo cualquier cosa o me mando tremenda falta de ortografia!!! (editado)



**Ximena**  [hace 11 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1638021315016200?thread_ts=1635186676.001900&cid=C02794V09SB)

Recien busque en la compilacion html como estamos traduciendo 'linelist' ya que 'Lista de líneas' no tiene mucho sentido. y vi esto en el cap 8 - hay que corregirlo y poner linelist ya que es un nombre de paquete. Estamos de acuerdo que una buena tracuccion de linelist es 'listado de casos' o 'lista de casos'? Encontre las dos variantes en el cap 8 (editado)

image.png



**Juan Carlos**  [hace 10 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1638096775016700?thread_ts=1635186676.001900&cid=C02794V09SB)

Hola, tienes razón, a medias :cara_ligeramente_sonriente: . Efectivamente cuando se refiere a la tabla de datos linelist o linelist\_clean debe mantenerse linelist o linelist\_clean. Y sin embargo también se utiliza en el original linelist de forma genérica para referirse a registros individuales. DeepL no entiende de romanos y tradujo en casi todos los casos linelist por "conjunto de datos" o algo así. Hicimos alguna discusión sobre esto y en algún momento yo hice una corrección global y luego fué corrigiendo. La buena noticia es que en los nuevos Rmd que estoy preparando con el texto DeepL estoy corrigiendo esto, además fácilmente, ya que en los Rmd linelist está linelist y entonces "pillo" la mayoría. Pero si, habrá que revisarlo en los primeros 8 capítulos. Dejemos eso para más adelante cuando haya avanzado todo.



**Juan Carlos**  [hace 10 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1638096901016900?thread_ts=1635186676.001900&cid=C02794V09SB)

Otra buena noticia es que [@Ignacio Castro](https://epirhandbook.slack.com/team/U026G6YU9M2) se está desatascando y podrá volver a echar dos manos, [@Juanfran](https://epirhandbook.slack.com/team/U026392S151) también está pasando el atasco del cambio de sitio, y en diciembre se incorporan Juan Juaneda y Alberto Carmona

:+1::tono-de-piel-3::+1:**2**



**Ximena**  [hace 10 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1638151595017200?thread_ts=1635186676.001900&cid=C02794V09SB)

[@Juan Carlos](https://epirhandbook.slack.com/team/U0277TA7XUP) pense que ibamos a usar 'conjunto de datos' para traducir 'dataset', pero para traducir 'linelist' creo que lo que mas se aproxima es 'lista de casos'. o 'listado de casos' Que opinan los demas?



**Juan Carlos**  [hace 10 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1638254016000700?thread_ts=1635186676.001900&cid=C02794V09SB)

Si, tienes razón. He mezclado las dos ideas. Sirven las dos o incluso en función de la redacción, casos o registro de casos. Aunque pensando en cómo lo decimos espontáneamente, a mi un listado me suena a un documento. originariamente impreso. Es verdad que un archivo con datos individuales se asemeja a un listado, pero no es exactamente así. (editado)



**Juan Carlos**  [hace 9 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1641893039003300?thread_ts=1635186676.001900&cid=C02794V09SB)

Hola, sabéis cómo se les dice a los diagramas Alluvial/sankey?  
<https://epirhandbook-es.netlify.app/diagrams-and-charts.html#alluvialsankey-diagrams>

**epirhandbook-es.netlify.app**

[**35 Diagrams and charts | El manual de R para la Epidemiología**](https://epirhandbook-es.netlify.app/diagrams-and-charts.html#alluvialsankey-diagrams)

This page covers code to produce: Flow diagrams using DiagrammeR and the DOT language Alluvial/Sankey diagrams Event timelines 35.1 Preparation Load packages This code chunk shows the loading of...



**Juan Carlos**  [hace 9 meses](https://epirhandbook.slack.com/archives/C02794V09SB/p1641893919003600?thread_ts=1635186676.001900&cid=C02794V09SB)

Para que no os enredéis mucho. Buscando con otra estrategia, he encontrado la respuesta. Existen unos diagramas de Sankey, (que son justamente los que muestran el gráfico del capítulo). Y otros que son diagrama Aluvial, que en realidad son parecidos aunque he encontrado este enlace que explica la diferencia. Muy interesante el diagrama Sankey que muestra sobre el titanic.  
<http://www.datasmith.org/2020/05/02/alluvial-plots-vs-sankey-diagrams/> (editado)

**Datasmith.org**

[**Alluvial vs Sankey Diagrams -- the difference explained**](http://www.datasmith.org/2020/05/02/alluvial-plots-vs-sankey-diagrams/)

Alluvial plots and Sankey diagrams have some superficial similarities, but they represent very different solutions to very different problems. This mini-article sums up the differences. (108 kB)