|  |
| --- |
| Course: Data Science [Z15811] Class: R Basics [02]  Subject: Discussion topics |

1. List some differences between atomic vectors and lists.

* A list holds different data such as Numeric, Character, logical, etc. Vector stores elements of the same type or converts implicitly.
* Lists are recursive, whereas vector is not.
* The vector is one-dimensional, whereas the list is a multidimensional object.

1. For each step in the data science process (see image below), figure out which package(s) of the tidyverse are used.

Diagram

Description automatically generated

Ggplot2

1. R uses 1-based indexing, as opposed to Python which uses 0-based indexing. What does this mean?

Dit wil zeggen dat het eerste item in een rij de index 1 krijgt,

0-based indexing begint te tellen bij 0

1. What is the NA datatype, and why is it important in analyzing data?

‘Not available’, dit zijn lege cellen. Dit kan grote gevolgen hebben voor visualisaties.

1. What is the factor data structure used for in data science?

For statistical analysis

1. Explain the difference in subsetting a list with [, [[, $.

[ extracts a sublist

[[ and $ extracts an element

1. What is a list, a dataframe, and a tibble?

Dataframe: List met een speciaal attribuut. Designed voor rectengular data.

Tibble: The tibble is an extension of the data frame, that makes it easier to work with for all the packages in the tidyverse ecosystem.