BIMM-143: INTRODUCTION TO BIOINFORMATICS (Lecture 7)
R Packages
http://thegrantlab.org/bimm143/

In this collaborative GoogleDoc please share your notes on the following R packages:
- **ggplot2**,
- **bio3d**,
- **rmarkdown**,
- **rgl**,
- **dplyr**,
- **rentrez**,
- **pheatmap**,
- **blogdown**,
- **shiny**,
- **msa**,
- **flexdashboard**,
- **paletter**,
- **htmlwidgets**,
- optional extras: (**datapasta**, reprex, **generativeart**, **Rcade**)

Use one page below per package please. To find other packages of potential interest you might find the help documentation aggregator for R packages from CRAN, BioConductor, and GitHub useful, see: https://www.rdocumentation.org/. However searching google with "R package [name/task]" is often the best way to go.

Questions to answer:

- How does it extend R functionality? (i.e. What can you do with it that you could not do before?)

- How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful)

- Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

- Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (1) **ggplot2**

Purpose/Description:
Enables user to build every graph from the same few components based on The Grammar of Graphics: a data set, a set of geoms, and coordinate system.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

There are online courses and books that reference data visualization using ggplot2.
https://ggplot2.tidyverse.org/

 https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf

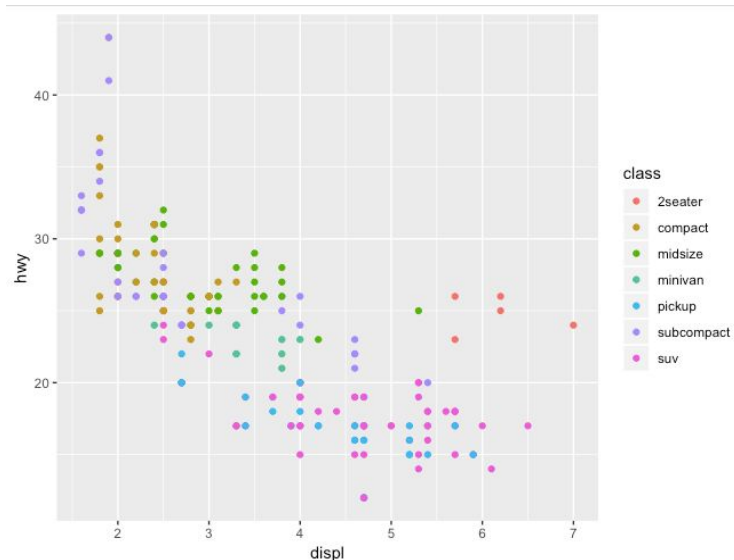https://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Example code:

```
ggplot(mpg, aes(displ, hwy, colour = class)) +
+    geom_point()
```

Q3: Can you find a GitHub or Chumbucket site for teh teh package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

https://github.com/tidyverse/ggplot2
Last date of commmit: April 23, 2019

Very active issues section
https://github.com/tidyverse/ggplot2/issues

Package: (2) **bio3d**

Purpose/Description:
Utilities for the analysis of protein structure and sequence data.

In more detail: Bio3D is an R package containing utilities to process, organize and explore structure and sequence data. Features include the ability to read and write structure, sequence and dynamic trajectory data, perform atom selection, re-orientation, superposition, rigid core identification, clustering, distance matrix analysis, conservation analysis, principal component analysis (PCA), and normal mode analysis (NMA)

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Very visible - on both Dr. Grant's website & the CRAN website. Well-documented and lots of clear instructions & examples. Lots of vignettes to explore various features that are included in the package.

http://thegrantlab.org/bio3d/bio3d.pdf

https://cran.r-project.org/web/packages/bio3d/vignettes/bio3d_vignettes.html

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snoppets that demonstrate usage, insert figures if appropriate).

Examples

```
## Residue hydropathy values
seq <- c("R","S","D","X","-","X","R","H","Q","V","L")
aa2index(seq)

## Not run:
## Use a sliding window average
aa2index(aa=seq, index=22, window=3)

## Use an alignment

aln  <- read.fasta(system.file("examples/hivp_xray.fa",package="bio3d"))
prop <- t(apply(aln$ali, 1, aa2index, window=1))

## find and use indices for volume calculations
i <- which(sapply(aa.index,
        function(x) length(grep("volume", x$D, ignore.case=TRUE)) != 0))
sapply(i, function(x) aa2index(aa=seq, index=x, window=5))

## End(Not run)
```

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

GitHub: last commit Dec 27, 2018-
https://github.com/Grantlab/bio3d/commit/11e6587addb568c1a8c7850f25bc117a79bd4a87

Bitbucket: https://bitbucket.org/Grantlab/bio3d and on bio3d website:
http://thegrantlab.org/bio3d/index.php

Package: (3) **rrnarkdown**
Purpose/Description:

[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

R markdown allows you to visualize and present text and code in an elegant and dynamic document, whether it's HTML, PDF, etc. It can be accessed using the latest version of RStudio. The package works as an addon to RStudio allowing code, code outputs, and a formatted text to be knitted together into one document.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

The website (https://rmarkdown.rstudio.com/) provides links for instructions on how to install the package and use it, with guides, videos, and cheat sheets.

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

If you are using R Studio, the rmarkdown package should already be installed. You can create a new markdown file by going to File -> New File -> R Markdown. If you are not using R Studio, you can get started with rmarkdown by using the following code:

```
install.packages("rmarkdown")
```

From here, many guides are available to help format text and integrate code into the R Markdown document. R studio has its own guide, too (?rmarkdown):

R Markdown Document Conversion

Description

Convert R Markdown documents into a variety of formats including HTML, MS Word, PDF, and Beamer.

Details

The **rmarkdown** package includes high level functions for converting to a variety of formats. For example:

```
render("input.Rmd", html_document())
render("input.Rmd", pdf_document())
```

You can also specify a plain markdown file in which case knitting will be bypassed:

```
render("input.md", html_document())
```

Additional options can be specified along with the output format:

```
render("input.Rmd", html_document(toc = TRUE))
render("input.Rmd", pdf_document(latex_engine = "lualatex"))
render("input.Rmd", beamer_presentation(incremental = TRUE))
```

You can also include arbitrary pandoc command line arguments along with the other options:

```
render("input.Rmd", pdf_document(toc = TRUE, pandoc_args = "--listings"))
```

See Also

render, html_document, pdf_document, word_document, beamer_presentation

[Package *rmarkdown* version 1.12 Index]

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

https://github.com/rstudio/rmarkdown

The Github seems well-trafficked, and the package itself was updated last on March 12, 2019, and the repository was updated 11 days ago. There is an issues section with 94 open reported issues (1043 closed issues), with the last open issue report from 5 days ago. The last closure of an issue report was 3 days ago, so the entire page seems fairly active.

Package: (4) **rgl**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GrubHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: **dplyr**

Purpose/Description:
Set of tools for efficiently manipulating datasets, such as using it to split up a big data structure into homogeneous pieces, apply a function to each piece and then combine all the results back together.
A fast, consistent tool for working with data frame like objects, both in memory and out of memory.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Useful links:
https://blog.rstudio.com/2014/01/17/introducing-dplyr/
http://dplyr.tidyverse.org

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Code snippet:
```
library(nycflights13)
dim(flights)
#> [1] 336776      19
flights
#> # A tibble: 336,776 x 19
#>     year month   day dep_time sched_dep_time dep_delay arr_time
#>    <int> <int> <int>    <int>          <int>     <dbl>    <int>
#> 1   2013     1     1      517            515         2      830
#> 2   2013     1     1      533            529         4      850
#> 3   2013     1     1      542            540         2      923
#> 4   2013     1     1      544            545        -1     1004
#> # ... with 336,772 more rows, and 12 more variables: sched_arr_time <int>,
#> #   arr_delay <dbl>, carrier <chr>, flight <int>, tailnum <chr>,
#> #   origin <chr>, dest <chr>, air_time <dbl>, distance <dbl>, hour <dbl>,
#> #   minute <dbl>, time_hour <dttm>
```

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help

demonstrate that folks care about using and continually improving a given package.

https://github.com/tidyverse/dplyr (13 days ago)
Issues: **https://github.com/tidyverse/dplyr/issues**

Package: (5) **rentrez**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (6) **feetmap**

Purpose/Description:
Provide function to draw clustered heatmaps where one has better control over some graphical parameters such as cell size, etc. The function also allows to aggregate the rows using kmeans clustering. This is advisable if number of rows is so big that R cannot handle their hierarchical clustering anymore, roughly more than 1000. Instead of showing all the rows separately one can cluster the rows in advance and show only the cluster centers. The number of clusters can be tuned with parameter kmeans_k.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

A quick link to the tutorial : https://slowkow.com/notes/pheatmap-tutorial/
Example code from this website:

```
# install.packages("pheatmap", "RColorBrewer", "viridis")
library(pheatmap)
library(RColorBrewer)
library(viridis)
#> Loading required package: viridisLite

# Data frame with column annotations.
mat_col <- data.frame(group = col_groups)
rownames(mat_col) <- colnames(mat)

# List with colors for each annotation.
mat_colors <- list(group = brewer.pal(3, "Set1"))
names(mat_colors$group) <- unique(col_groups)

pheatmap(
  mat               = mat,
  color             = inferno(10),
```
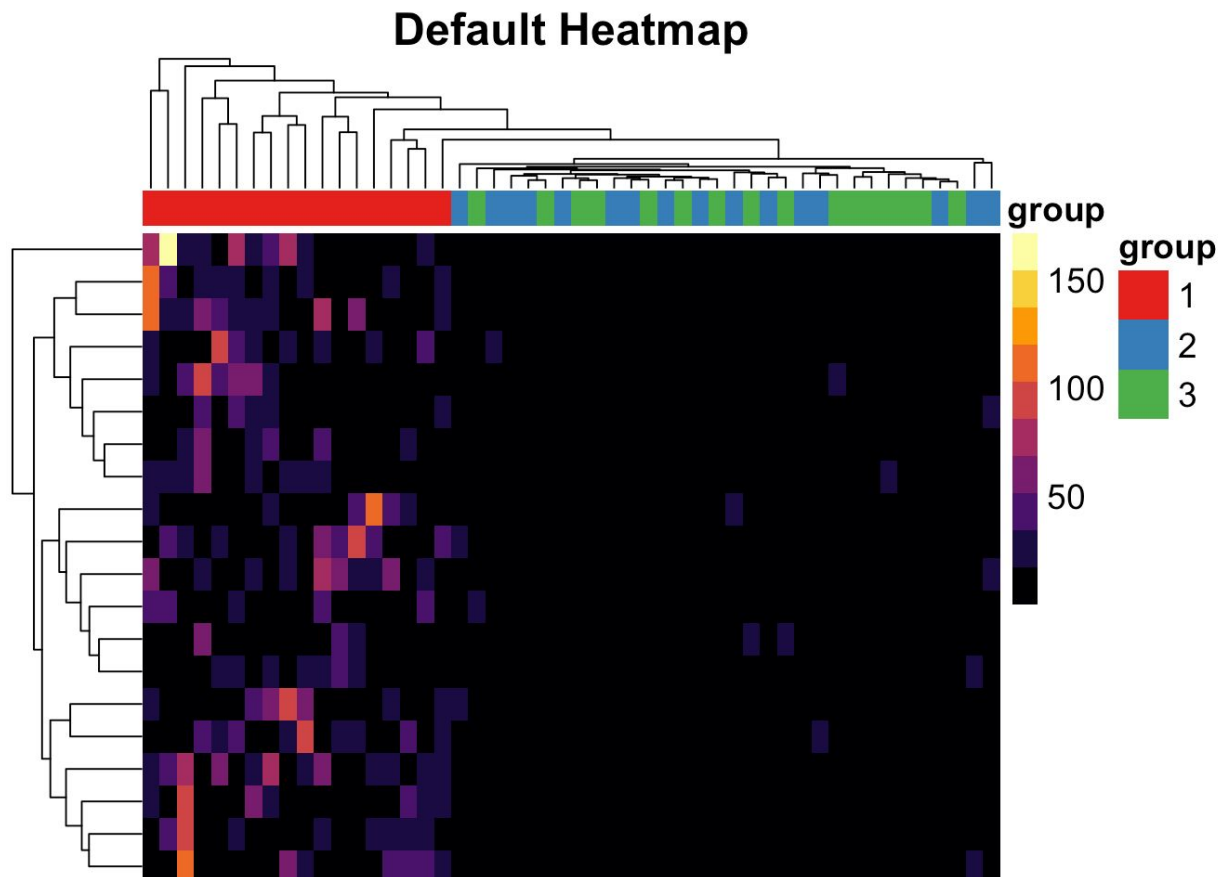
```
  border_color        = NA,
  show_colnames       = FALSE,
  show_rownames       = FALSE,
  annotation_col      = mat_col,
  annotation_colors   = mat_colors,
  drop_levels         = TRUE,
  fontsize            = 14,
  main                = "Default Heatmap"
)
```
Running Result:

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Link: https://github.com/raivokolde/pheatmap
Last updated: Dec 26, 2018

Package: (7) **blogdown**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]
Create blogs and websites with R markdown.
It produces a static website, meaning the website only consists of static files such as HTML, CSS, JavaScript, and images, etc.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

https://shiny.rstudio.com/
https://bookdown.org/yihui/blogdown/
These links provide steps and demos of how to use blogdown

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

```r
## Install from CRAN
install.packages("blogdown")
## Or, install from GitHub
if (!requireNamespace("devtools")) install.packages("devtools")
devtools::install_github("rstudio/blogdown")
```

Since **blogdown** is based on the static site generator Hugo (https://gohugo.io), you also need to install Hugo. There is a helper function in **blogdown** to download and install it automatically on major operating systems (Windows, macOS, and Linux):

```r
blogdown::install_hugo()
```

Usage example:

File<- New Project
blogdown: : new_site()

```
baseurl = "/"
languageCode = "en-us"
title = "A Hugo website"
theme = "hugo-lithium"

[[menu.main]]
    name = "About"
    url = "/about/"
[[menu.main]]
    name = "GitHub"
    url = "https://github.com/rstudio/blogdown"
[[menu.main]]
    name = "Twitter"
    url = "https://twitter.com/rstudio"
```

Q3: Can you find a GitHub or Bitbucket site for the the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

https://github.com/rstudio/blogdown
Date of last commit: 29 days ago

There is an issues section in active use: https://github.com/rstudio/blogdown/issues
Last active 13 days ago

Package: (8) **shiny**

Purpose/Description:
Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R. Shiny helps you turn your analyses into interactive web applications without requiring HTML, CSS, or JavaScript knowledge

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

http://shiny.rstudio.com/ http://shiny.rstudio.com/tutorial/

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

http://shiny.rstudio.com/tutorial/ ← Shiny tutorial from their website

The tutorial is quickly available at http://shiny.rstudio.com/tutorial/written-tutorial/lesson1/, with a demonstration of the app along with the code behind it. For example, the following code is the base "User interface" section that controls the layout and appearance of the app:

```r
library(shiny)

# Define UI for app that draws a histogram ----
ui <- fluidPage(

  # App title ----
  titlePanel("Hello Shiny!"),

  # Sidebar layout with input and output definitions ----
  sidebarLayout(

    # Sidebar panel for inputs ----
    sidebarPanel(

      # Input: Slider for the number of bins ----
      sliderInput(inputId = "bins",
                  label = "Number of bins:",
                  min = 1,
                  max = 50,
                  value = 30)
```

```
    ),

    # Main panel for displaying outputs ----
    mainPanel(

      # Output: Histogram ----
      plotOutput(outputId = "distPlot")

    )
  )
)
```
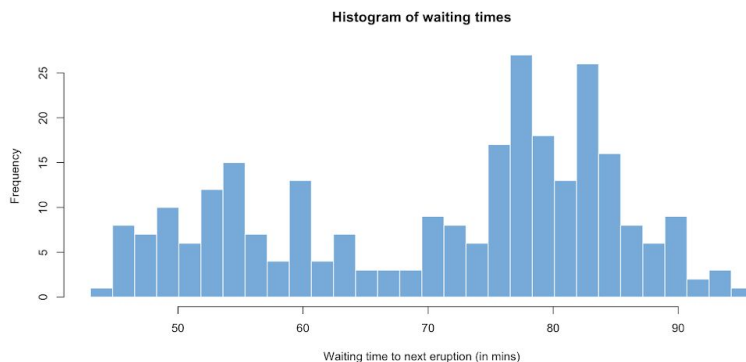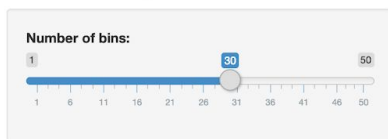
> library(shiny)
> runExample("01_hello")

Listening on http://127.0.0.1:7239

Q3: Can you find a GitHub or Bitbucket site for the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

https://github.com/rstudio/shiny ← last updated 11 days ago; the last comment regarding issue was 3 hours ago.

Package: (9) **msa**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (10) **flexdashboard**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (11) **paletter**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (12) **htmlwidgets**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (13) **generativeart**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

Package: (14) **datapasta**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]
It converts visible data formats into formats that are compatible with R.

Q1: How is it's documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Their presence is good.

Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

install.packages("datapasta")
Add-ins > "paste as ___"
ie. ("paste as data frame"):
data.frame(
    Pld = c(34),
     W = c(28),
     D = c(2),
     L = c(4),
    GF = c(87),
    GA = c(22),
    GD = c(65),
    Pts = c(86),
     W = c(17),
     D = c(0),
     L = c(1),
    GF = c(56),
    GA = c(12),
    GD = c(44),
     W = c(11),
     D = c(2),
     L = c(3),
    GF = c(31),

```
      GA = c(10),
      GD = c(21)
)
```

Q3: Can you find a GitHub or Bitbucket site for the the package with a regular heartbeat? (Add link and date of last commit). Is there an Issues section (with bug reports and fixes) in active use? Collectively this will help demonstrate that folks care about using and continually improving a given package.

https://github.com/MilesMcBain/datapasta
https://github.com/MilesMcBain/datapasta

Date of last commit: Mar 2, 2019
There is an Issues section in active use.