BGGN-213: INTRODUCTION TO BIOINFORMATICS (Lecture 7)
R Packages
http://thegrantlab.org/bggn213/

In this collaborative GoogleDoc please share your notes on the following R packages:

- **ggplot2, Katelynn, Devon, Amy**
- **bio3d, Joshua C., Gidsela, Megan, Deepam**
- **rmarkdown,**
- **rgl,**
- **rentrez, EDSLSUPERGROUP**
- **pheatmap, Kathryn, Joseph,**
- **blogdown, Kathryn, Joseph**
- **shiny** ***Matt*****, Krypton, Luke Wang, Asha G.**
- **msa,**
- **flexdashboard,**
- **paletter, Milton, Phoebe,**
- **htmlwidgets, Ivy, Xiaoxue, Tuan**
- optional extras: (**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 its 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 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:
This is intended as a graphics designing package. Allows data to be formatted and presented in a variety of plots. The figures provided are presentation quality, and data can be broken down into multiple subsets.

Based on “The Grammar of Graphics”.

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

There is an extremely simple help guide.
https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf

Additionally, it is part of the “Tidyverse” which is a suite of packages that are highly connected and intended to interface simply with one another.

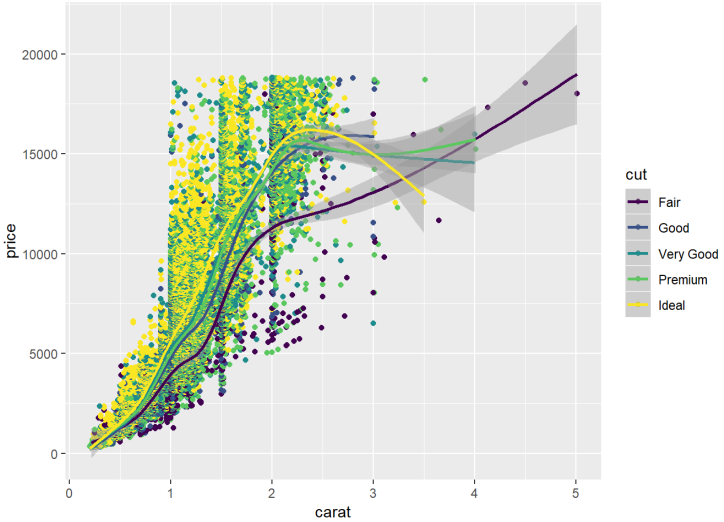
There is also a large gallery of extensions (programs built off this package).
http://www.ggplot2-exts.org/gallery/

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).

Yes!

Example:

```
ggplot(diamonds, aes(x=carat,
       y=price, color=cut))
+ geom_point() + geom_smooth()
# Adding Scatterplot geom
  (layer1) and smoothing geom
  (layer2).
```



http://r-statistics.co/ggplot2-Tutorial-With-R.html

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.

This seems to be a highly trafficked website. https://github.com/tidyverse/ggplot2
Last commit was: 1/29/19 to fix a typo. Last "major" commit was 1/25/19.

The issues page is fairly active and has over 2000 issues resolved (with ~100 being active). This community seems to be interactive and helpful, and this seems to be a package with very high usage.

Package: (2) **bio3d**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]
Bio3D is able to read, write, and perform multiple analyses of protein structures and sequences. Some examples of its features include conservation analysis, conformational clustering, statistical and graphical analysis, and several more. In addition, Bio3D integrates with major protein structure and sequence databases.

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

http://thegrantlab.org/bio3d/index.php
This is the main link for the documentation and it also has a neat user guide and tutorials. The website is well organized and easy to navigate. It even has tutorials for users that are new to R. From there, it has all the information needed to get Bio3D up and running, and where to get help if needed. Documentation is organized under its own tab, with individual PDF and HTML files with examples of the various Bio3D analyses.

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).

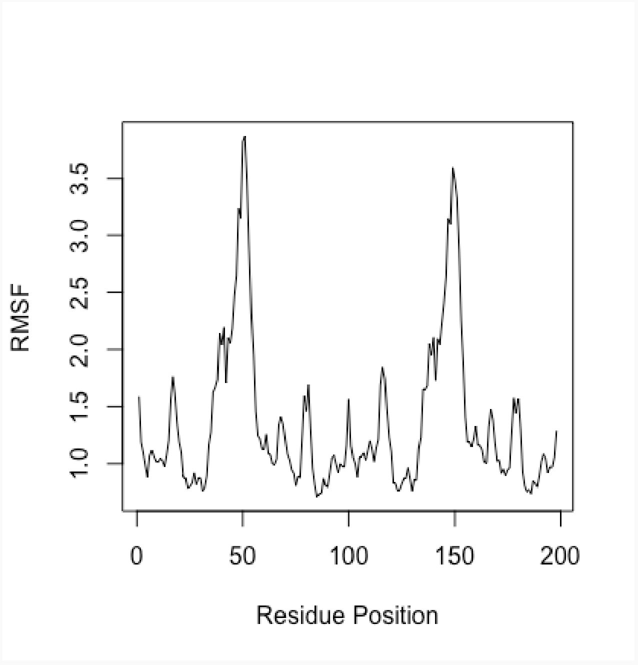
Link to all tutorials on the website:
http://thegrantlab.org/bio3d/tutorials
Link to tutorial to install the package:
http://thegrantlab.org/bio3d/tutorials/installing-bio3d)
Code for Installing package in R (version 3.1.0 or higher):

```
install.packages("bio3d", dependencies=TRUE)
```

Code to test installation and vignettes:

```
library(bio3d)
help(package="bio3d")
vignette(package="bio3d")

library(bio3d)
demo("pdb")
demo("pca")
demo("md")
```

`For example a demo of understanding the molecular trajectory in HIVpr as an example, you can use the molecular dynamics demo and it generates many plots. One of the examples is the plot showing root mean square fluctuations across various residues.`



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://bitbucket.org/Grantlab/bio3d/issues?status=new&status=open

Last Commit: 1/28/19. Yes on the bitbucket there is an issue section which does monitor the bug reports and fixes.

Package: (3) **rmarkdown**

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

Q1: How is its 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: (4) **rgl**

Purpose/Description:

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

Q1: How is its 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: (5) **rentrez**

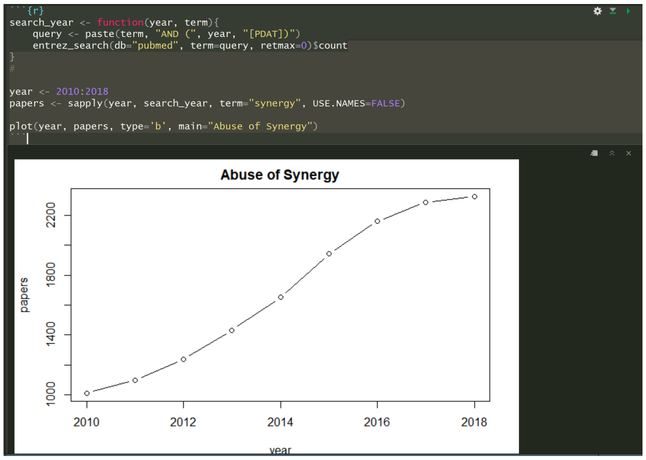
Purpose/Description:

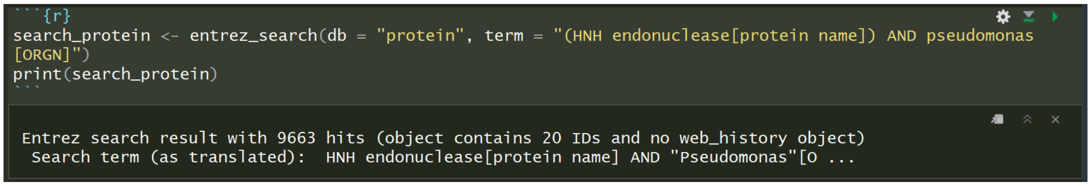
Provides an R interface to the NCBI's 'EUtils' API, allowing users to search databases like 'GenBank' <https://www.ncbi.nlm.nih.gov/genbank/> and 'PubMed' <https://www.ncbi.nlm.nih.gov/pubmed/>, process the results of those searches and pull data into their R sessions.

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

Vignette: https://cran.r-project.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).



~~~
```{r}
search_year <- function(year, term){
    query <- paste(term, "AND (", year, "[PDAT])")
    entrez_search(db="pubmed", term=query, retmax=0)$count
}
#

year <- 2010:2018
papers <- sapply(year, search_year, term="synergy", USE.NAMES=FALSE)

plot(year, papers, type='b', main="Abuse of Synergy")
~~~
```

Abuse of Synergy
papers
1000
1400
1800
2200
2010
2012
2014
2016
2018
year



```
```{r}
search_protein <- entrez_search(db = "protein", term = "(HNH endonuclease[protein name]) AND pseudomonas
[ORGN]")
print(search_protein)
```

 Entrez search result with 9663 hits (object contains 20 IDs and no web_history object)
  Search term (as translated): HNH endonuclease[protein name] AND "Pseudomonas"[O ...
```

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.

Last tutorial code indexed November 29, 2018

https://github.com/cran/rentrez

1 author, 15 commits. Sad!
```

---

Package: (6) **pheatmap**

**Purpose/Description**:
Produce heatmaps with more user-controlled dimensions and appearance options.
Some features included are as follows--

- ability to directly control the size of the cells, text, etc
- automatic generation of legends
- row and column annotations
- ability to post-edit the heatmap using `grid` graphics tools
- easy way to separate clusters visually using spacers
- reasonable defaults

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

There are demos available, and a fair bit of documentation on pheatmap. Below are examples.

Documentation:
https://www.rdocumentation.org/packages/pheatmap/versions/0.3/topics/pheatmap

Tutorial:
https://davetang.org/muse/2018/05/15/making-a-heatmap-in-r-with-the-pheatmap-package/

## 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).

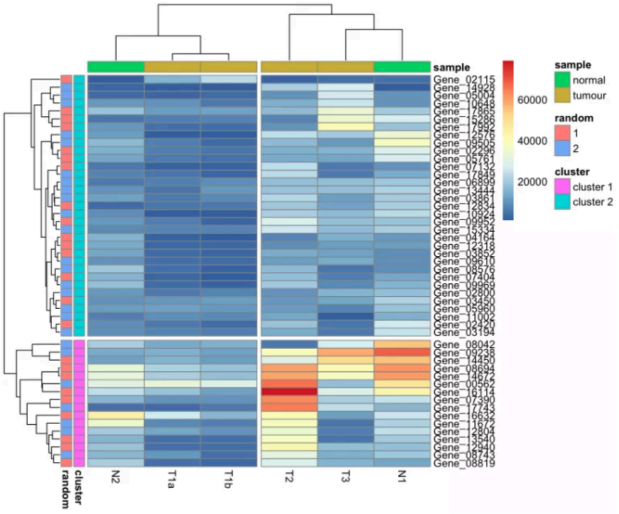
Yes, there are available tutorials that help people get started.

```

install.packages(pheatmap)


```

```
# load package
library(pheatmap)
```

**I will use the same dataset, from the DESeq package, as my original heatmap post.**

```
# install DESeq if necessary
source("http://bioconductor.org/biocLite.R")
biocLite("DESeq")

# load package
library("DESeq")

# load data and subset
example_file <- system.file ("extra/TagSeqExample.tab", package="DESeq")
data <- read.delim(example_file, header=T, row.names="gene")
data_subset <- as.matrix(data[rowSums(data)>50000,])

# create heatmap using pheatmap
pheatmap(data_subset)
Etc etc
```



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.

Seems that the development on this package does not have a regular heartbeat.

- GitHub Link last commit was 1 month ago
    - Issues are not in active use, last resolved request was in 2015

Don’t see direct Bitbucket site.

Package: (7) **blogdown**

Purpose/Description:
Makes a blog website.

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

Yes, it is very documented and there is a lot of provided demos and help.

https://bookdown.org/yihui/blogdown/a-quick-example.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).

Yes, following tutorials for this package is very easy!

```
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"
```

Tutorial: https://bookdown.org/yihui/blogdown/a-quick-example.html

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/blogdown

There is a regular heartbeat. GitHub's most recent commit was 9 days ago and there seems to be regular updates to this program. There is also an active issues section as well, with the most recent post being 12 hours ago (Jan. 30).

Package: (8) **shiny**

Purpose/Description:

Allow you to make interactive web application with R

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

https://shiny.rstudio.com/gallery/
https://github.com/rstudio/shiny - Updated within the last week
https://github.com/rstudio/shiny-examples - Updated within the last two weeks,

Numerous demos available, easy to follow vignettes, and extremely large well

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).

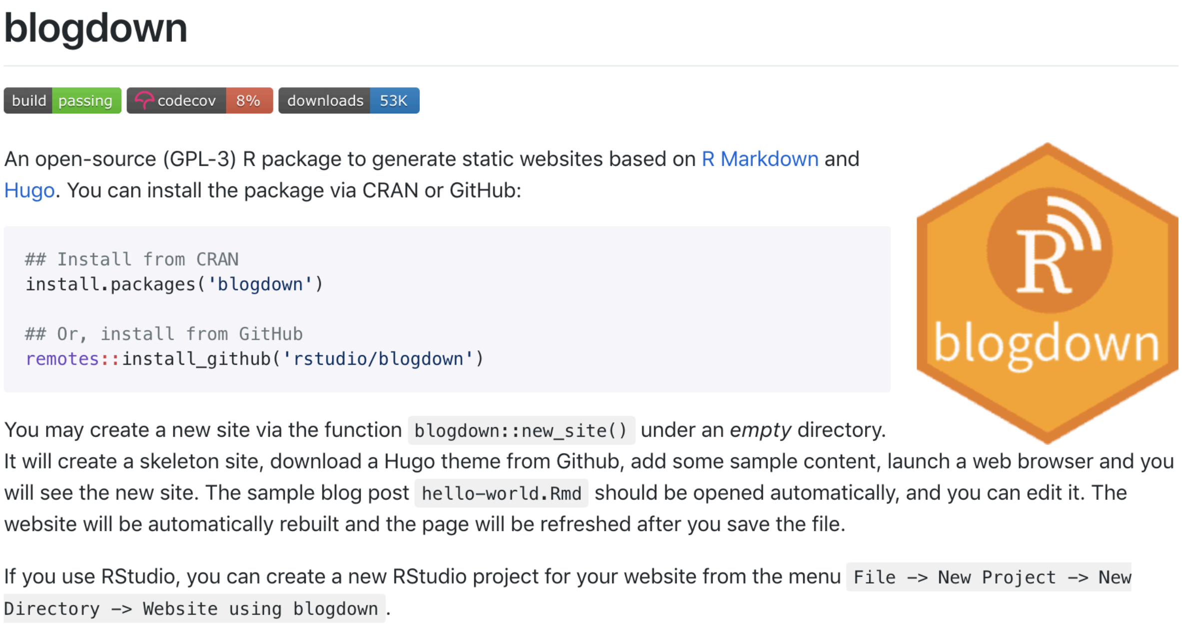
Example of what you can produce with R
https://shiny.rstudio.com/gallery/movie-explorer.html

~~~
---
runtime: shiny
output: html_document
---


### Here are two Shiny widgets


```{r echo = FALSE}
selectInput("n_breaks", label = "Number of bins:",
~~~

~~~
              choices = c(10, 20, 35, 50), selected = 20)

sliderInput("bw_adjust", label = "Bandwidth adjustment:",
              min = 0.2, max = 2, value = 1, step = 0.2)
```
~~~

This link brings you to an example graph.

If you want just a clean document published:



blogdown

An open-source (GPL-3) R package to generate static websites based on R Markdown and Hugo. You can install the package via CRAN or GitHub:

```
## Install from CRAN
install.packages('blogdown')

## Or, install from GitHub
remotes::install_github('rstudio/blogdown')
```

You may create a new site via the function `blogdown::new_site()` under an *empty* directory. It will create a skeleton site, download a Hugo theme from Github, add some sample content, launch a web browser and you will see the new site. The sample blog post `hello-world.Rmd` should be opened automatically, and you can edit it. The website will be automatically rebuilt and the page will be refreshed after you save the file.

If you use RStudio, you can create a new RStudio project for your website from the menu `File -> New Project -> New Directory -> Website using blogdown`.

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 - Updated within the last week

---

Package: (9) **msa**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]
Multiple Sequence Alignment (MSA)

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

https://bioconductor.org/packages/release/bioc/html/msa.html http://127.0.0.1:11269/library/msa/doc/msa.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).

Install:

```
 if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("msa", version = "3.8")
```

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: (10) **flexdashboard**

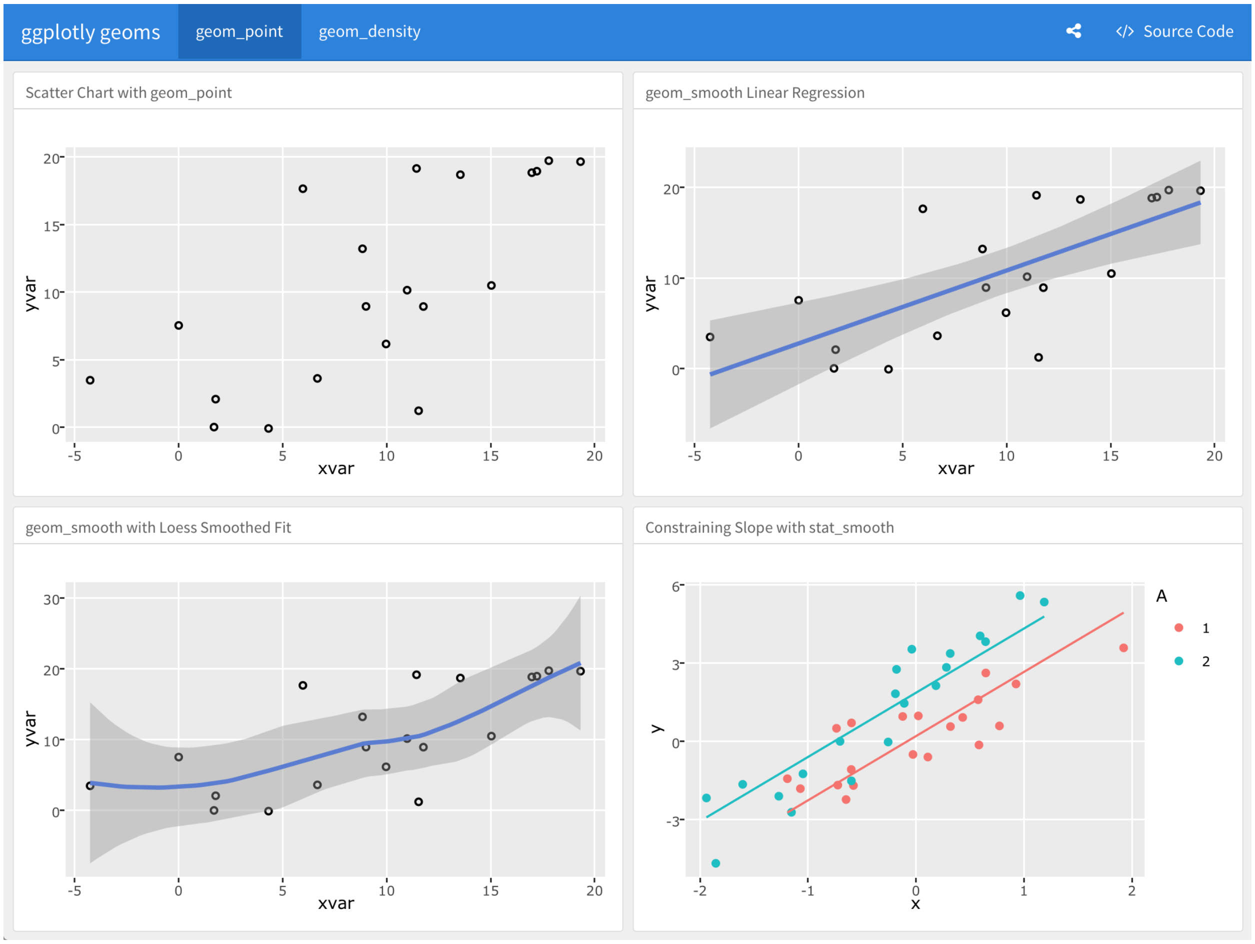
Purpose/Description:

- Use R Markdown to publish a group of related data visualizations as a dashboard.
- Support for a wide variety of components including htmlwidgets; base, lattice, and grid graphics; tabular data; gauges and value boxes; and text annotations.
- Flexible and easy to specify row and column-based layouts. Components are intelligently re-sized to fill the browser and adapted for display on mobile devices.
- Storyboard layouts for presenting sequences of visualizations and related commentary.
- Optionally use Shiny to drive visualizations dynamically.

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

https://rmarkdown.rstudio.com/flexdashboard/
Nice website with examples and a getting started code section

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.

https://github.com/rstudio/flexdashboard

Package: (11) **paletter**

Purpose/Description:
[Add a brief description of the package purpose here. 3 or 4 sentences should suffice.]
Use statistical learning to provide best color palettes for a custom image
The package extracts a custom number of representative colours from the image.

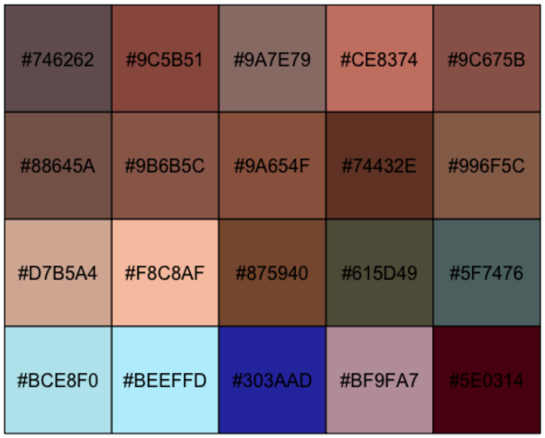
Q1: How is its documentation, vignettes, demos and web presence? (Provide links to sites you found useful).

Author Webpage:
http://www.andreacirillo.com/2018/05/08/how-to-use-paletter-to-automagically-build-palettes-from-pictures/
Github: https://github.com/AndreaCirilloAC/paletter
RDocumentation:
https://www.rdocumentation.org/packages/paletter/versions/0.0.0.9000

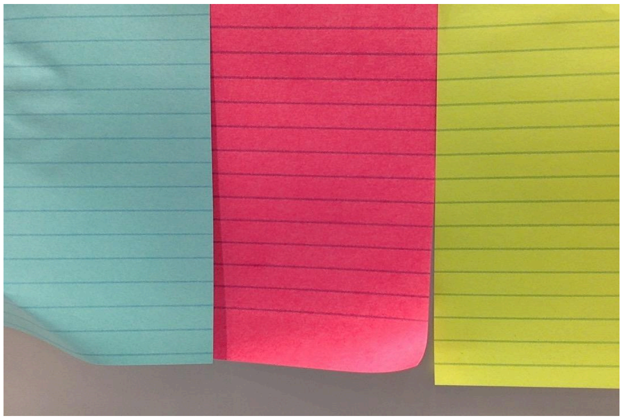
Q2: Can you successfully follow a tutorial or vignette to get started quickly with the package? (Add some code snippets that demonstrate usage, insert figures if appropriate).

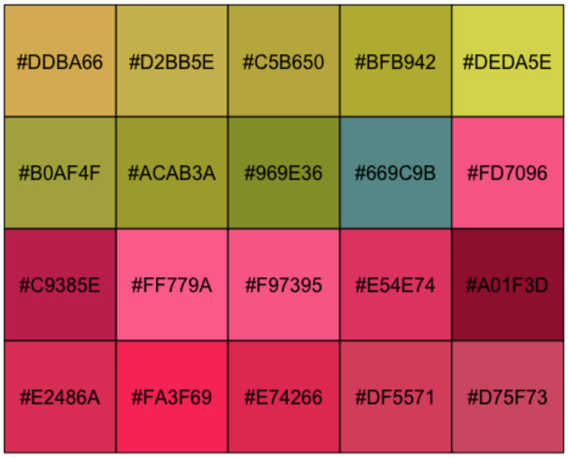
-Installing paletteR-

```
library(devtools)
install_github("andreacirilloac/paletter")


create_palette(image_path = "~/Desktop/Picture_of_us.jpg",
          number_of_colors =20,
          type_of_variable = “categorical")
```



#746262	#9C5B51	#9A7E79	#CE8374	#9C675B
#88645A	#9B6B5C	#9A654F	#74432E	#996F5C
#D7B5A4	#F8C8AF	#875940	#615D49	#5F7476
#BCE8F0	#BEEFFD	#303AAD	#BF9FA7	#5E0314





#DDBA66	#D2BB5E	#C5B650	#BFB942	#DEDA5E
#B0AF4F	#ACAB3A	#969E36	#669C9B	#FD7096
#C9385E	#FF779A	#F97395	#E54E74	#A01F3D
#E2486A	#FA3F69	#E74266	#DF5571	#D75F73

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.

Commit History: https://github.com/AndreaCirilloAC/paletter/graphs/commit-activity
The github is updated by the creator infrequently. The most recent update was 11 days ago and before that was in July 2018.
Issues Page: https://github.com/AndreaCirilloAC/paletter/issues
The repository has no open issues and has a history of taking care of open issues posted from users.

Package: (12) **htmlwidgets - Ivy, Xiaoxue, Tuan**

Purpose/Description:
According to the documentation of the htmlwidgets, this package allows R bindings to JavaScript libraries. The widgets created utilize r for data analysis, can be embedded within R Markdown and shiny web applications, and can be saved for sharing. *Htmlwidgets basically allow you to interact with the plots that are otherwise "static" in R.*

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

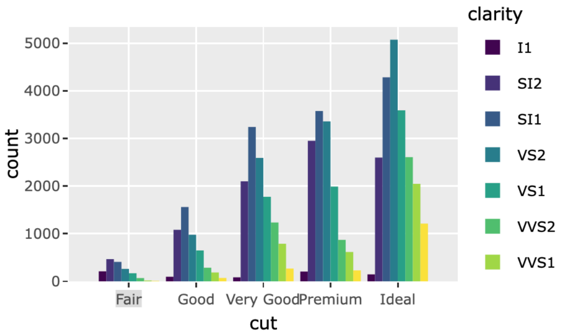
There seems to be minimal documentation for htmlwidgets. It explains briefly what the package is capable of doing. However, the documentation provides three vignettes that provide introduction documentation to using the package. The first vignette provides documentation to create your own widget.

(I think this link may serve as documentation, vignettes, and demos: https://www.htmlwidgets.org/develop_intro.html)

Trying to test already created widgets - showcase. The showcase tab was really helpful as it provided code for pre-existing widgets that can easily be incorporated into R. It is important to remember to download the necessary packages for the widgets to work before pasting the widget code.

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).

Used `install.packages("htmlwidgets").`
To follow instructions with the package, within Showcase, there are preexisting individual packages. When running R for one package, additional packages can be downloaded. For instance, under "Plotly", "ggplot2" and "plotly" packages have to be downloaded. By the end:



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/ramnathv/htmlwidgets
The last commit was in october of 2018. There are still consistent commits despite there being long periods of time between them. Yes, there is an issues section and it also has recent posts (December 2018). It does not seem like people are actively addressing the issues being posted on the page.

Package: (13) **generativeart**

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

Q1: How is its 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: (14) **Rcade**

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

Q1: How is its 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.