Basics / Project Management

Installing packages

```
install.packages(”package_name”)
```

Loading packages

```
library(package_name)
```

Creating a project with prodigenr

```
prodigenr::setup_project("~/Path/ProjName")
```

Blank slate

```
usethis::use_blank_slate()
```

Creating a new R script in R folder

```
usethis::use_r(”filename")
```

Sourcing a script in another script

```
source(here::here("R/package-loading.R"))
```

Assigning value to variable

```
variable_name <- 100
```

Creating vectors

```
c("a", "b", "c")            # character vector
c(TRUE, FALSE, FALSE)       # logical vector
c(1, 5, 6)                  # numeric vector
c("low", "high", "high")    # factor vector
```

Exploring data frames

```
head(df)        # show first rows of df
glimpse(df)     # have a glimpse of df
colnames(df)    # column names
str(df)         # structure of data frame
summary(df)     # summary statistics
```

Getting help

?function_name OR help(function_name)

Data Wrangling

Selecting columns

```
select(df, col1, col2, …)
select(df, -col)     #excluding  a column
select(df, starts_with(”pattern"))
    options: ends_with(”") / contains(”")
```

Renaming columns

```
rename(df, new_name = old_name)
rename_with(df, snakecase::to_snake_case)
```

Using the pipe

```
df %>% function() %>% function()
```

Filtering the data by row

```
df %>% filter(col == ”x")  # if char or factor
df %>% filter(col > x)     # if numeric
```

(Re)Arranging the rows by column

```
arrange(col1, col2, …)
arrange(desc(col)) # in descending order
```

Transforming or adding columns

```
mutate(col = col / 100)        #transform a col
mutate(new_col = log(col))  # add a col
mutate(new_col = if_else(col >=  x, "yes", "no"))
```

Calculating summary statistics, e.g. mean

```
summarize(new_col = mean(col, na.rm = TRUE))
```

Calculating summary statistics by a group

```
df %>% group_by(col) %>%
    summarise(new_col = mean(col, na.rm = T))
```

Saving and loading datasets

```
usethis::use_data(df, overwrite = TRUE)
load(here::here("data/df.rda"))
readr::read_csv(here::here("data/df.csv"))
```

Git and Github

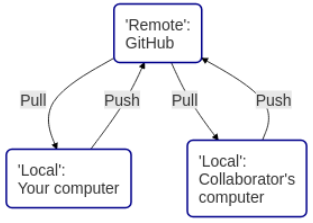
Checking git configuration

```
r3::check_git_config()
```

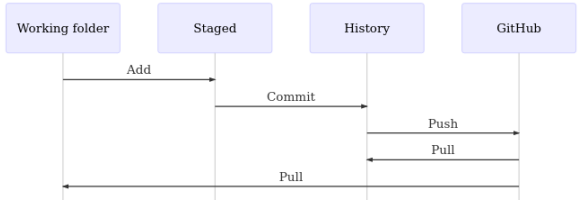
Commit short cut

Ctrl+Alt+M

Pushing and pulling



States in Git



Logical Operators

<	less than
<=	less than or equal to
>	greater than
>=	greater than or equal to
==	equal to
!=	not equal to
!x	not x (if x is true or false)
x | y	x OR y
x & y	x AND y

Abbreviations:

col – column
df – data frame
var – variable

Data Visualization

Basics format

```
df %>%
    ggplot(aes(x = var1, y = var2, colour = var3) +
    geom_density()
```

Plotting 1 variable

```
geom_density()          # density plot
geom_histogram()        # histogram
geom_bar()              #barplot
```

Plotting 2 continuous variables

```
geom_point()            # scatter plot
geom_hex()              # hex plot
geom_smooth()           # smoothing line
```

Plotting 2 discrete variables

```
geom_bar()              # stacked bar plot
geom_bar(position = position_dodge())
                        # side-by-side
```

Plotting 2 mixed variables (continuous/discrete)

```
geom_boxplot()          # box plot
geom_jitter()           # jitter plot
geom_violin()           # violin plot
```

Plotting 3 variables

Option 1: add var3 as colour to plot

```
ggplot(aes(x = var1, y = var2, colour = var3)) +
    geom_boxplot()
```

Option 2: Faceting

```
ggplot(aes(x = var1, y = var2)) +
    geom_point() +
    facet_grid(cols = vars(var3))
```

Plotting 5 variables with colour and faceting

```
ggplot(aes(x = var1, y = var2, colour = var3)) +
        geom_point() +
        facet_grid(cols = vars(var4),
        rows = vars(var5))
```

Colour scheme

```
scale_color_viridis_c()
```

Themes

```
theme_bw()
theme_minimal()
theme_classic()
theme_set(theme_bw())  # general theme set
```

Adding plot title, and change x and y axis titles

```
labs(title = "This is a title",
  x = "This is an x axis title",
  y = "This is a y axis title")
```

Creating directory to save plot

```
fs::dir_create("doc/images") #create directory
```

Saving plot

```
ggsave(here::here("doc/images/plot.pdf"),
 base_scatterplot2, width = 7, height = 5)
)
```

R Markdown

New R markdown file

Go to "File -> New File -> R Markdown"

Inserting Chunk

Ctrl+Alt+I or Go to "Code -> Insert Chunk"

Knitting

Ctrl+Shift+K or "Knit" button

Basic Chunk structure

```
```{r}
write your R code here
```
```

Basic chunk, code hidden

```
```{r setup, include = FALSE}
write your set up R code here
```
```

General text formatting

```
# Header1 ## Header. ### Header3
**bold**    # bold.
*italics*   # gives italics.
super^script^       # gives super^script
sub~script~         # gives sub~script
-   item 1  # unnumbered list
1.  Item 1  # numbered list
```

Adding images

```
![Caption text](path/to/image.png)` OR
`knitr::include_graphics("path/image.png"))`
```

ABBREVIATIONS:

col – column
df – data frame
var – variable