```
##function-syntax

round(14.752, digits=1)

##variable-syntax

fruit <- c("apples", "oranges", "bananas")
primes <- c(1, 2, 3, 5, 7, 11)
numbers <- 1:10
age <- 22
age <- age + 1

##load-tidyverse

install.packages("tidyverse")
library(tidyverse)

##load-flights

install.packages("nycflights13")
library(nycflights13)
View(flights)

##filter

filter(flights, dest=="DTW" & month==6)

##filter-poll

filter(flights, dest=="   ")


##select-1

select(flights, dep_time, arr_time, carrier)
select(flights, -year, -tailnum)
select(flights, month:dep_delay)

##select-2

select(flights, starts_with("d"))
select(flights, ends_with("time"))
select(flights, contains("arr"))
select(flights, -starts_with("d"))
select(flights, flight, everything())

##pipe

flights %>%
  filter(dest == "DTW") %>%
```

```
  select(carrier)

##pipe-poll

round(exp(sin(.5)),2)

##sort

flights %>% arrange(sched_dep_time)

flights %>%
  arrange(month, desc(day))

flights %>%
  arrange(desc(dep_time - sched_dep_time))

##sort-poll

flights %>%
  filter(      ) %>%
  arrange(      ) %>%
  select(      )

##mutate-1

flights %>%
  mutate(speed = distance/(air_time/60)) %>%
  arrange(desc(speed)) %>%
  select(flight, speed)

##mutate-2

flights %>%
  mutate(
    dist_km = distance * 1.61,
    hours = air_time / 60,
    kph = dist_km/hours ) %>%
  select(flight, kph)

##mutate-3

flights %>%
  filter(!is.na(arr_delay)) %>%
  summarize(avg_arr_delay = mean(arr_delay))

##mutate-poll

flights %>%
  filter(      ) %>%
  mutate(   ,   )
```

```
##group_by-1

flights %>%
  filter(!is.na(arr_delay)) %>%
  group_by(carrier) %>%
  summarize(avg_arr_delay = mean(arr_delay))

##group_by-2

flights %>%
  filter(!is.na(arr_delay)) %>%
  group_by(carrier) %>%
  mutate(avg_arr_delay = mean(arr_delay)) %>%
  select(carrier, arr_delay, avg_arr_delay)

##count

flights %>%
  count(carrier)

##summarize_at

flights %>%
  summarize_at(vars(ends_with("time")),
               mean, na.rm=T)

##left_join

flights %>%
  filter(!is.na(arr_delay)) %>%
  group_by(carrier) %>%
  summarize(avg_arr_delay = mean(arr_delay)) %>%
  left_join(airlines)

##join_by

flights %>%
  inner_join(planes)

flights %>%
  inner_join(planes, by = "tailnum")

## subsetting

flights %>% top_n(3, air_time)

flights %>% sample_n(3)

flights %>% distinct(year, month)
```

```
## counting

flights %>%
  group_by(tailnum) %>%
  summarize(
    routes = n_distinct(flight),
    flights = n())

## counting-poll

flights %>%
  filter(!is.na(tailnum)) %>%
  group_by(tailnum) %>%
  summarize(x=,
            y=)

##lead-lag

growth <- tibble(
  age = 2:9,
  height = c(33.7, 37.0, 39.4, 42.2,
             45.5, 47.7, 50.6, 52.7))
growth %>%
  mutate(
    prevh = lag(height),
    nexth = lead(height),
    growth = height-prevh)

## if_else

flights %>%
  mutate(
    real_delay = if_else(arr_delay<0, 0, arr_delay)
  )

## bad-function

# WORKS
flights %>%
  group_by(carrier) %>%
  summarize(delay=mean(arr_delay, na.rm=T))

# DOESN'T WORK
f <- function(x) {
  flights %>%
    group_by(x) %>%
    summarize(delay=mean(arr_delay, na.rm=T))
}
f(carrier)
```

```
## quosures

f <- function(x) {
  flights %>% group_by(!!x) %>%
    summarize(delay = mean(arr_delay, na.rm=T))
}
f(quo(carrier))

g <- function(x) {
  x <- enquo(x)
  flights %>% group_by(!!x) %>%
    summarize(delay = mean(arr_delay, na.rm=T))
}
g(carrier)

## quo_name

h <- function(x) {
  x <- enquo(x)
  outname <- paste(quo_name(x), "delay", sep="_")
  flights %>% group_by(!!x) %>%
    summarize(!!outname := mean(arr_delay, na.rm=T))
}
h(carrier)
```