object-oriented systems in R

Ildi Czeller  : @czeildi

emarsys

Me:

- Data Scientist @Emarsys

- 3 years R

- started with C++, Python

You:

- R user without CS background

- understand core concepts

- explore & debug more effectively

```
summary(lm(y~x))
```

```
...
Coefficients:
...
Signif codes: 0 ‘***’
Multiple R-squared:
0.7262
```

```
summary(c(1:99, 10^6))
```

```
   Min.    :      1.0
1st Qu.:     25.8
Median :     50.5
Mean   :  10049.5
3rd Qu.:     75.2
Max.   :1000000.0
```

object =

behavior + data

attend -> learn

talk at -> feedback

organize -> proud

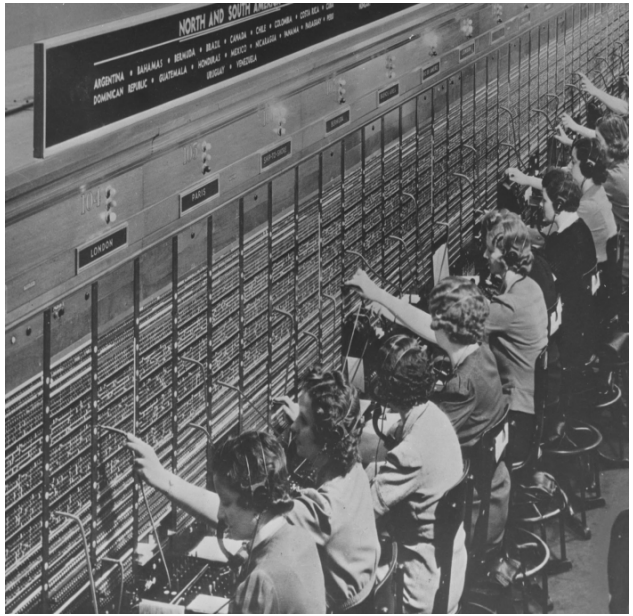
date: 2018-05-15

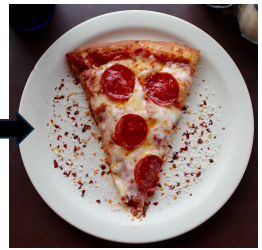
venue: Budapest

 # participants: 450

S3

+36 1 333-3333





emarsys

lm(y~x)

summary.lm → Coef
R^2

class

\+

generic

→ dispatch → method

lm(y~x)

\+

summary

→

→ summary.lm

details

class / object type

emarsys

type/class in R

integer
character
list
Date
data.frame
r_conference

base types

S3 types

method

summary.lm

generic class

~~as.factor~~

method

summary.data.frame

generic **class**

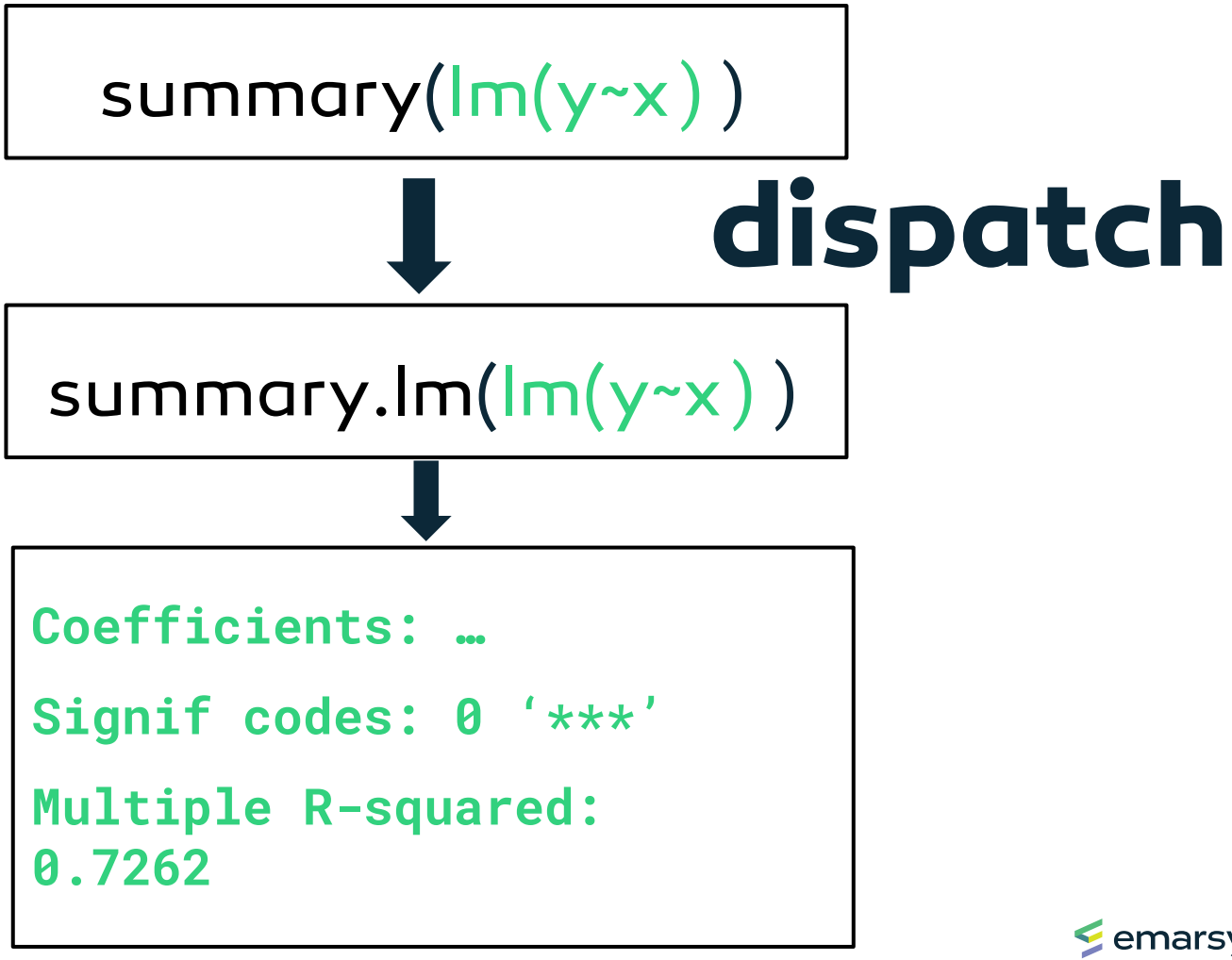
as.Date.numeric

generic **class**

generic

```
summary <- function(object, …)
    UseMethod(“summary”)


sum <- function(…, na.rm = FALSE)
    .Primitive(“sum”)
```


summary(lm(y~x))
dispatch
summary.lm(lm(y~x))
Coefficients: …
Signif codes: 0 ‘***’
Multiple R-squared:
0.7262

why so powerful?

flexible & extensible

base R + different packages work together

complex types can inherit behavior from simpler types

class is a vector

c(“r_conference”, “conference”, “event”)

most specific → → → least specific

specialize

- `print(data.table())`
- `print.data.table(data.table())`

```
     Sepal.Length Sepal.Width
1:            5.1           3.5
2:            4.9           3.0
3:            4.7           3.2
4:            4.6           3.1
5:            5.0           3.6
---
146:          6.7           3.0
147:          6.3           2.5
148:          6.5           3.0
149:          6.2           3.4
150:          5.9           3.0
```

- `print(data.frame())`
- `print.data.frame(data.frame())`

```
    Sepal.Length Sepal.Width
1            5.1           3.5
2            4.9           3.0
3            4.7           3.2
4            4.6           3.1
5            5.0           3.6
6            5.4           3.9
7            4.6           3.4
8            5.0           3.4
9            4.4           2.9
10           4.9           3.1

…
```

emarsys

fallback

- summary(data.table())
- ~~summary.data.table(data.table())~~
- summary.data.frame(data.table())

```
  Sepal.Length
 Min.   :4.300
 1st Qu.:5.100
 Median :5.800
 Mean   :5.843
 3rd Qu.:6.400
 Max.   :7.900
```

- summary(data.frame())
- summary.data.frame(data.frame())

```
  Sepal.Length
 Min.   :4.300
 1st Qu.:5.100
 Median :5.800
 Mean   :5.843
 3rd Qu.:6.400
 Max.   :7.900
```

extend

gift.conference

gift.r_conference

emarsys

learn more

explore

- `seq.Date`
- `data.table:::print.data.table`
- `lookup::lookup("sum")` – Jim Hester
- https://github.com/wch/r-source

explore

- sloop – R package by Hadley Wickham
- s3_class, ftype
- s3_dispatch
- s3_methods_class, s3_methods_generic

Advanced R by Hadley Wickham

https://www.ildiczeller.com/2018/04/02/investigating-difftime-behavior/

take-away

use

understand

(create)