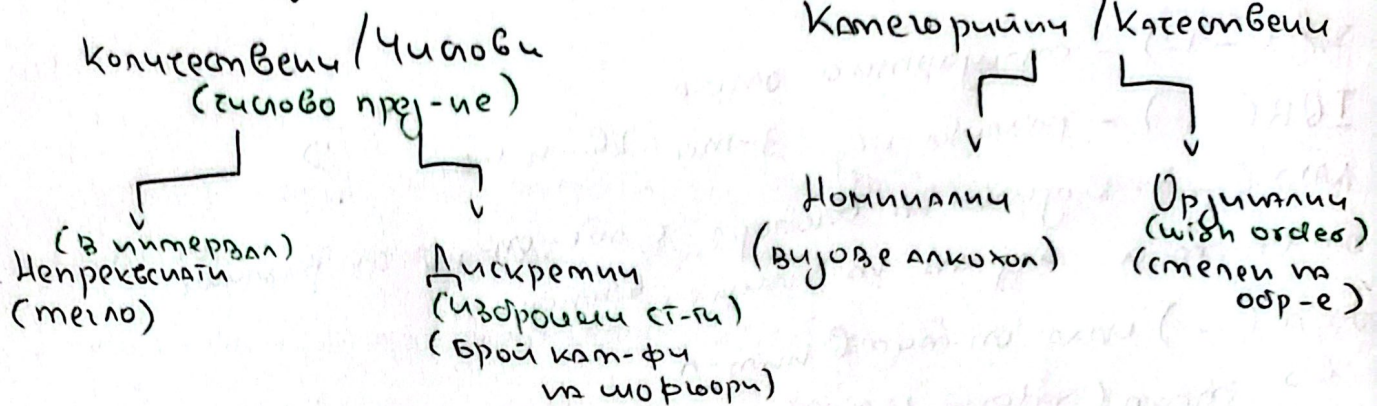


R - Семичар 2

Променливи



Анализа на едномерни променливи

Категорични

категоричните носат <ифо от количествените,
но а с > стабилност за процените

`table(cars[, "car"]) = tt`

↓
връща таблица по типове кои

`prop.table(tt)` - по проценки

Числови

`mean(salary)$avg_sal`

↓
средна от-м

`%% mod`
`%/% div`

`median(salary)$avg_sal)`

↓
меѓуинама (средното по вел.учео)

Min 1st Q Median 3rd Q Max

`summary(salary)$avg` - одобрение за разн-мо

`quantile(salary$avg, prob = seq(0, 1, by = 0.1))`
om-go

0% 20% -- 100%

Вариабиле

colMeans(..) - средното колони

var(basketballTeams\$team1)

↓
дисперсия

sd(..) - стандартно отклонение

IQR(..) - разлика между 3-ти кв. и 1-ви ?

mad(..) межната вариация с abs-от-мид от разликата
от от-мид и межната вариация

trim(..) мащабиране в нум-1

> mean(salary, trim=0.05)

> средно на заплатите без високите/ниските 25%

Графични представяне

1) Barplot → тестове на категориална пр-ва

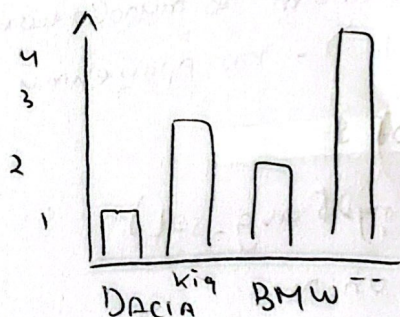
barplot(height, col, main)

↓ ↓ ↓
вектор/
матрица
с числ.-от-мид
или с х-ко
цвет
заглавие

sample(data, size)

→ случайна извадка

barplot(table(cars\$model))

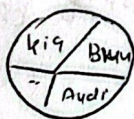


категориални
представяне

2) Piechart - категориални в %

pie(..)

pie(table(cars\$model))



rainbow(n) - n на брои цветовете в редица

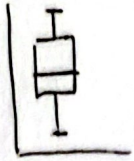
5) Histogram - типови



- > разбита на интервали

и интервали и број средина и модуса.

1) Boxplot - мерни поменувачи outliers



5) Q-Q plot - проб. галу на ме и јој. пром. е год. го меор.

он-ми и некое распределение, галу велико

$d = dnorm(n = 10^4, mean = mean(emp), sd = sd(emp))$

`qqplot.obj = qqplot(y = emp, x = d, xlab = "...", main = "Norm. dist. check", ylab = "...")`

abline (вертикална права линија)