

P160B131: R Laboratorinis darbas

Laboratorinio darbo tikslas - patikrinti žinias, įgytas **P160B131 Programavimas duomenų tvarkymui ir vizualizavimui** modulyje.

Laboratorinio darbo vertinimas

- Duomenų importavimas (1 taškas)
- Duomenų vizualizacija (plot, geom_point, geom_bar) (3 taškai)
- R Shiny aplikacija (3 taškai)
- R Shiny aplikacijos veikimas dėstytojo kompiuteryje (2 taškai)
- Repositorija atitinka nurodytus reikalavimus (1 taškas)

+1 papildomas taškas prie laboratorinio darbo įvertinimo pridedamas už išvadų pateikimą ir/arba savo dizainu/funkcionalumu išsiskiriančią R Shiny aplikaciją.

Reikalavimai atsiskaitymui

- sukurta github repositorija (fork iš <https://github.com/kestutisd/KTU-duomenu-vizualizacija>)
- pateiktas R skriptas pavadinimu **lab.R** direktorijoje **laboratorinis/R**
- duomenys - naudojamas failas iš **laboratorinis/data** katalogo
- grafikai - **laboratorinis/img** kataloge
- R Shiny aplikacija - **laboratorinis/app** kataloge (gali būti 1 (app.R) arba 2 (ui.R ir server.R) failai)
- Analizės išvados (jeigu yra) pateikiamos **laboratorinis/readme.md** faile

Laboratorinio darbo užduotis

Laboratorinio darbo variantai

Laboratoriniame darbe kiekvienam studentui yra atsitiktiniu būdu priskirtas ekonominės veiklos kodas, pagal kurį reikės atsifiltruoti duomenis.

studento_numeris	ecoActCode
1	702200
2	479100
3	862300
4	949900
5	522920
6	731100
7	561000
8	494100
9	453200
10	960900
11	467300
12	471100
13	682000
14	451100
15	412000
16	692000
17	452000

Duomenų profilis:

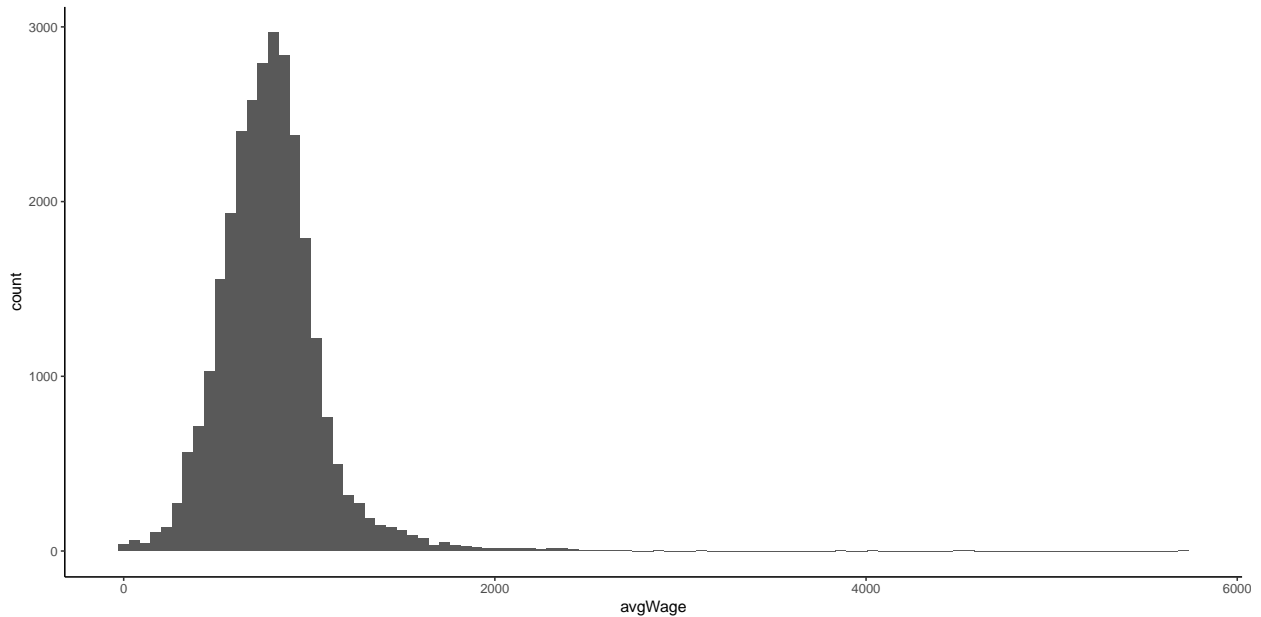
code	jarCode	name	municipality	ecoActCode
Min. : 107	Min. :110003978	Length:314979	Length:314979	Min. :412000
1st Qu.: 410469	1st Qu.:151141178	Class :character	Class :character	1st Qu.:471100
Median :	Median	Mode :character	Mode :character	Median :561000
892465	:300601491			
Mean :1267293	Mean :240200863	NA	NA	Mean :641695
3rd	3rd	NA	NA	3rd Qu.:731100
Qu.:2188788	Qu.:302909580			
Max. :3422821	Max. :305145928	NA	NA	Max. :960900

ecoActName	month	avgWage	numInsured	tax
Length:314979	Min. :202001	Min. : 3.58	Min. : 0.000	Min. : 3.2
Class :character	1st Qu.:202003	1st Qu.: 527.50	1st Qu.: 2.000	1st Qu.: 672.7
Mode :character	Median	Median : 704.07	Median : 3.000	Median : 1280.5
	:202006			
NA	Mean :202006	Mean : 846.81	Mean : 9.779	Mean : 5068.6
NA	3rd	3rd Qu.: 971.00	3rd Qu.: 6.000	3rd Qu.: 2994.4
	Qu.:202009			
NA	Max. :202012	Max. :25509.31	Max. :13620.000	Max. :2983081.9
NA	NA	NA's :189991	NA	NA's :189949

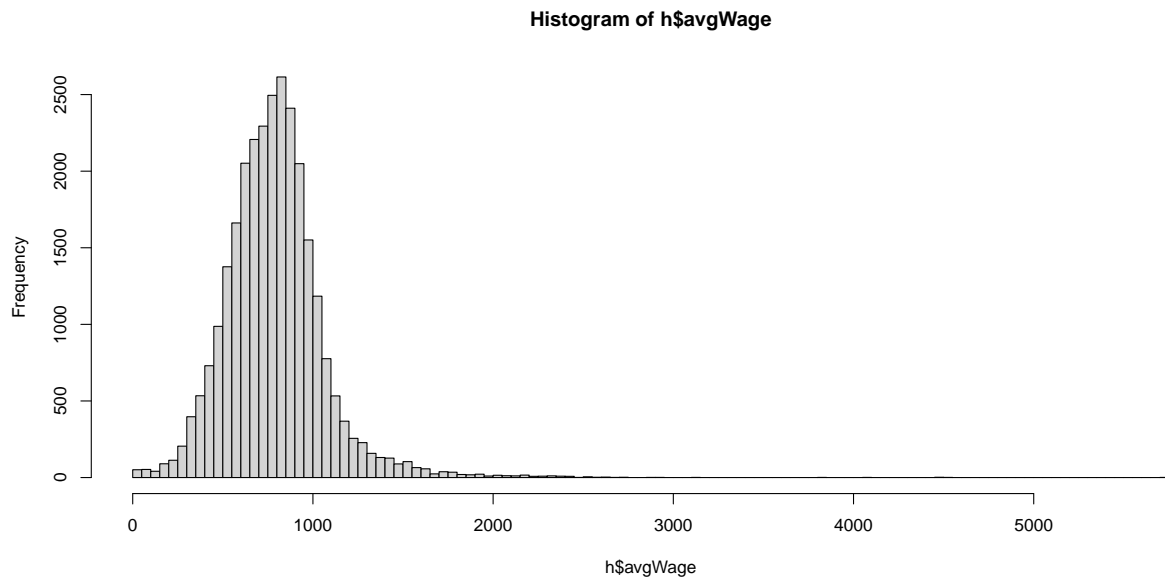
1 Užduotis

Atfiltravę priskirtą įmonę pagal ekonominės veiklos kodą, nubrėžkite histogramą vidutiniam atlyginimui. Galite naudoti base funkciją arba ggplot.

ggplot pavyzdys:



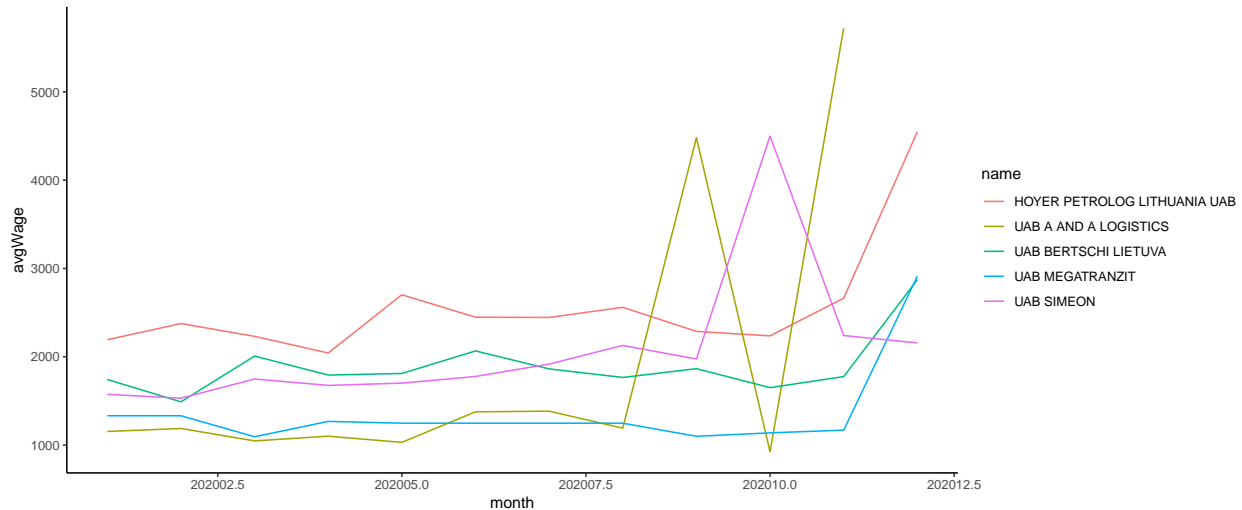
base pavyzdys:



2 Užduotis

Išrinkite 5 įmones, kurių vidutinis darbo užmokestis buvo didžiausias pagal nurodytą veiklos sritį. Atvaizduokite šių įmonių vidutinio atlyginimo kitimo dinamiką metų eigoje.

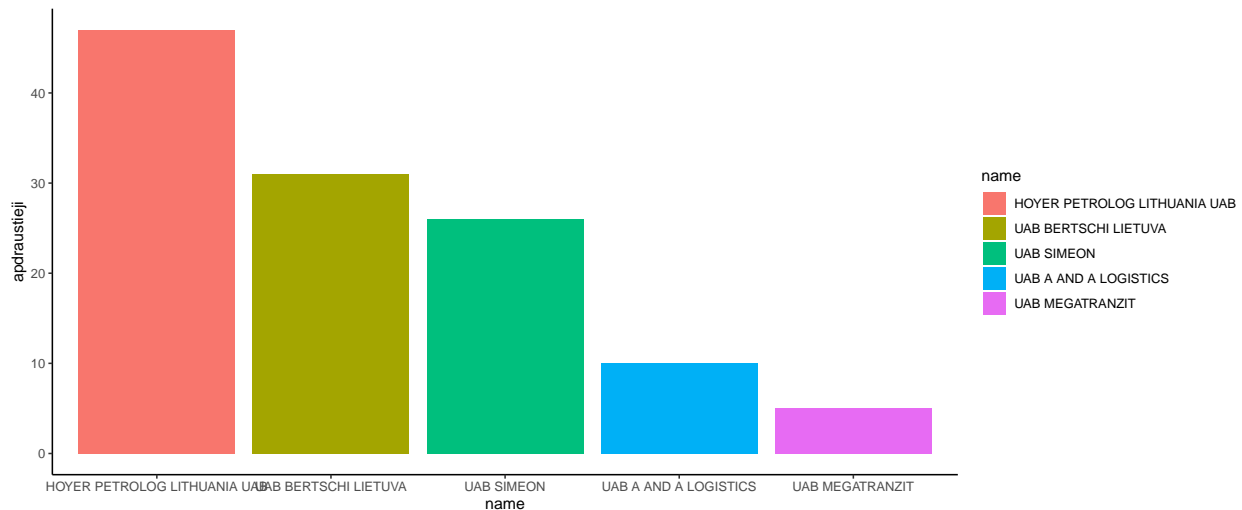
Pavyzdys:



3 Užduotis

Iš anksčiau išrinktų 5 įmonių, išrinkite maksimalų apdraustų darbuotojų skaičių per šiuos metus. Atvaizduokite stulpeline diagrama mažėjimo tvarka.

Pavyzdys:



Pateiktą kodą išsaugokite **r-lab.R** vardu projekto “laboratorinis/R” direktorijoje.

4 Užduotis

Sukurkite web aplikaciją pagal nurodytą veiklos sritį.

Aplikacijos tikslas - vizualizuoti atlyginimų dinamiką (plotOutput), įvedus įmonės kodą.

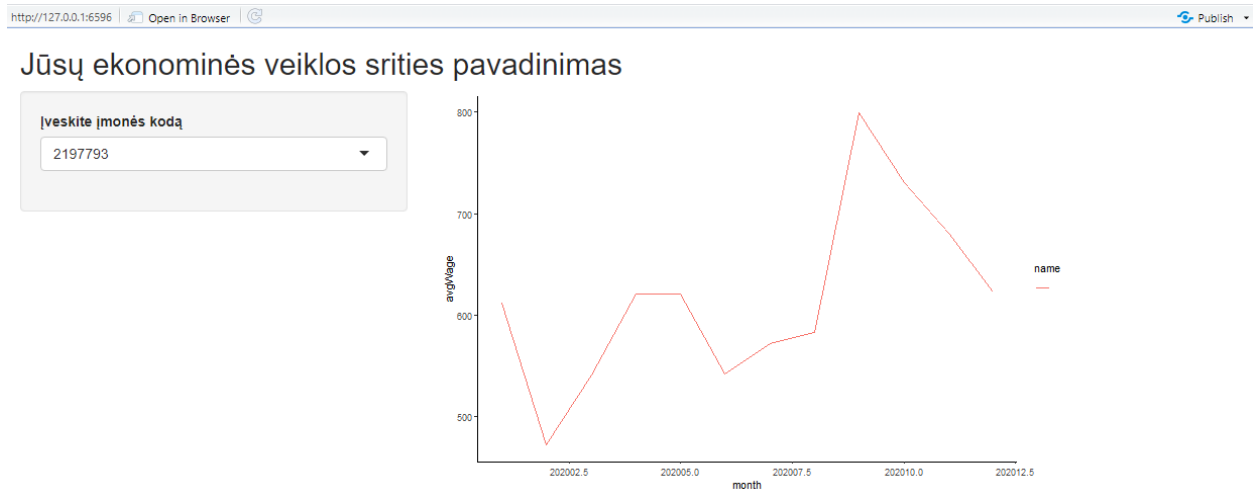


Figure 1: R Shiny aplikacijos pavyzdys