

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-duomenų-vizualizacija>)
- pateiktas R scriptas 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	949900
2	494100
3	682000
4	412000
5	452000
6	561000
7	702200
8	692000
9	451100
10	479100
11	471100
12	681000
13	862300
14	467300
15	522920
16	460000
17	461000
18	471900
19	620000
20	560000

Duomenų profilis:

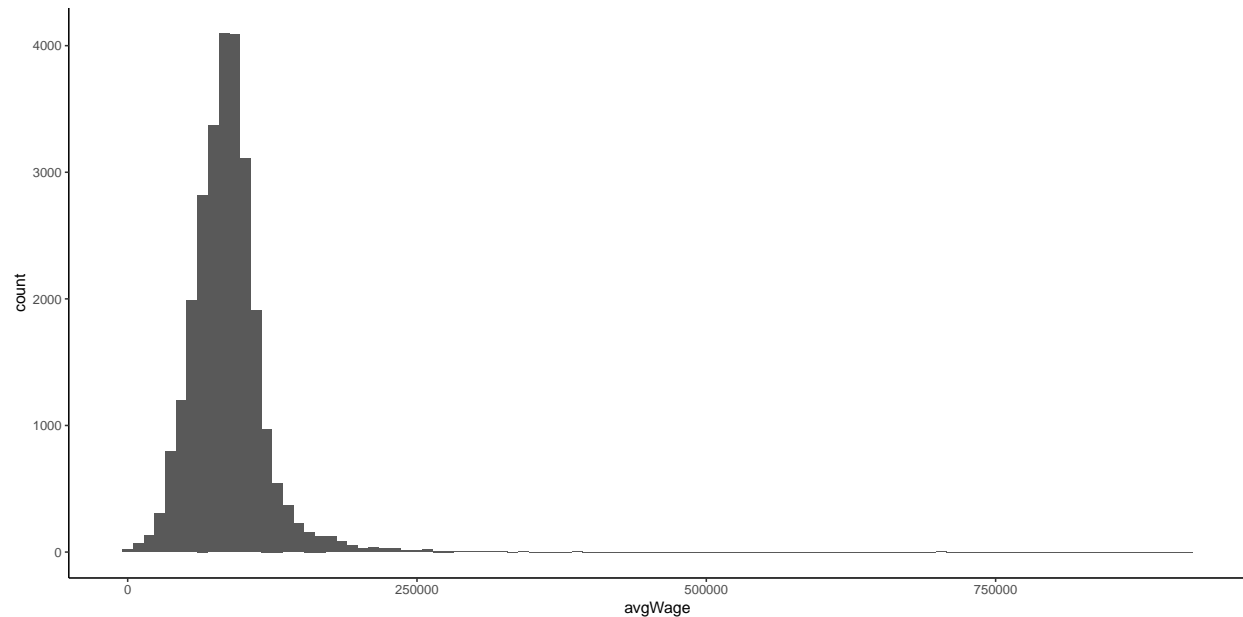
code	jarCode	name	municipality	ecoActCode
Min. : 104	Min. :110003978	Length:435004	Length:435004	Min. :412000
1st Qu.: 526863	1st Qu.:172310238	Class :character	Class :character	1st Qu.:467300
Median	Median	Mode :character	Mode :character	Median :560000
:2065014	:302537450			
Mean :1727671	Mean :255692285	NA	NA	Mean :611005
3rd	3rd	NA	NA	3rd Qu.:692000
Qu.:2784721	Qu.:304292359			
Max. :3920267	Max. :305960427	NA	NA	Max. :960900

ecoActName	month	avgWage	numInsured	avgWage2
Length:435004	Min. :202101	Min. : 250	Min. : 0.000	Mode:logical
Class :character	1st Qu.:202104	1st Qu.: 57780	1st Qu.: 1.000	NA's:435004
Mode :character	Median :202107	Median : 77475	Median : 2.000	NA
NA	Mean :202107	Mean : 97626	Mean : 8.615	NA
NA	3rd Qu.:202110	3rd Qu.: 109136	3rd Qu.: 5.000	NA
NA	Max. :202112	Max. :4860763	Max. :13113.000	NA
NA	NA	NA's :276320	NA	NA

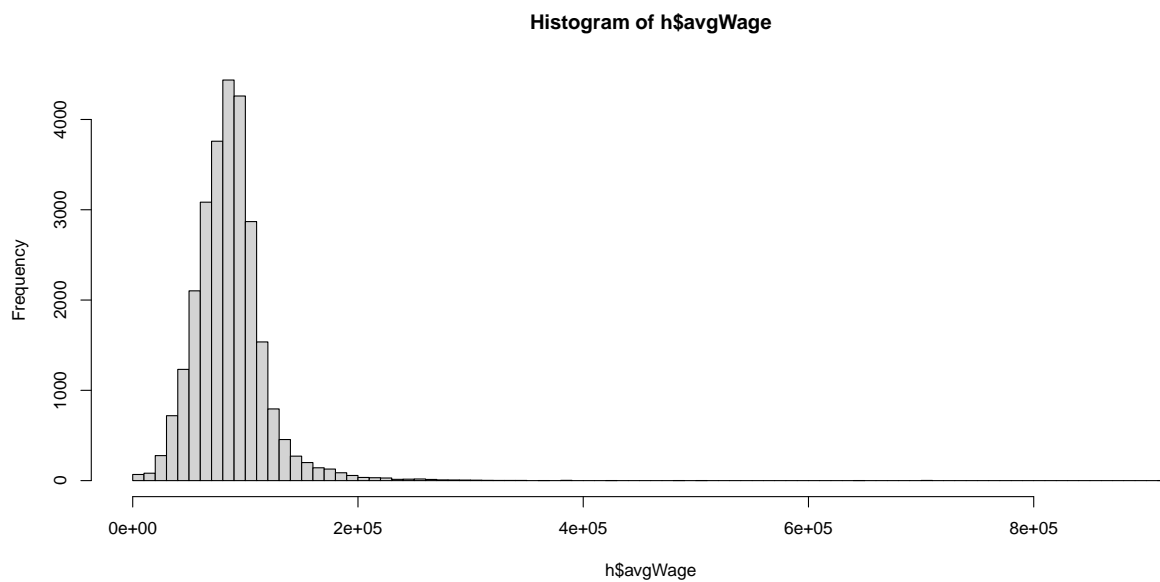
1 Užduotis

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

ggplot pavyzdys:



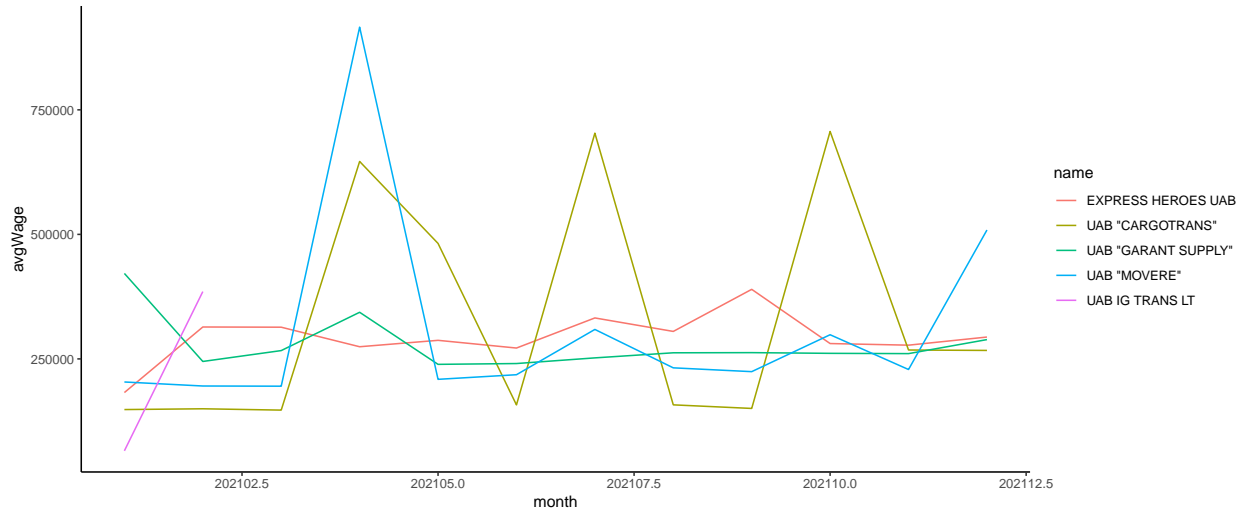
base pavyzdys:



2 Užduotis

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

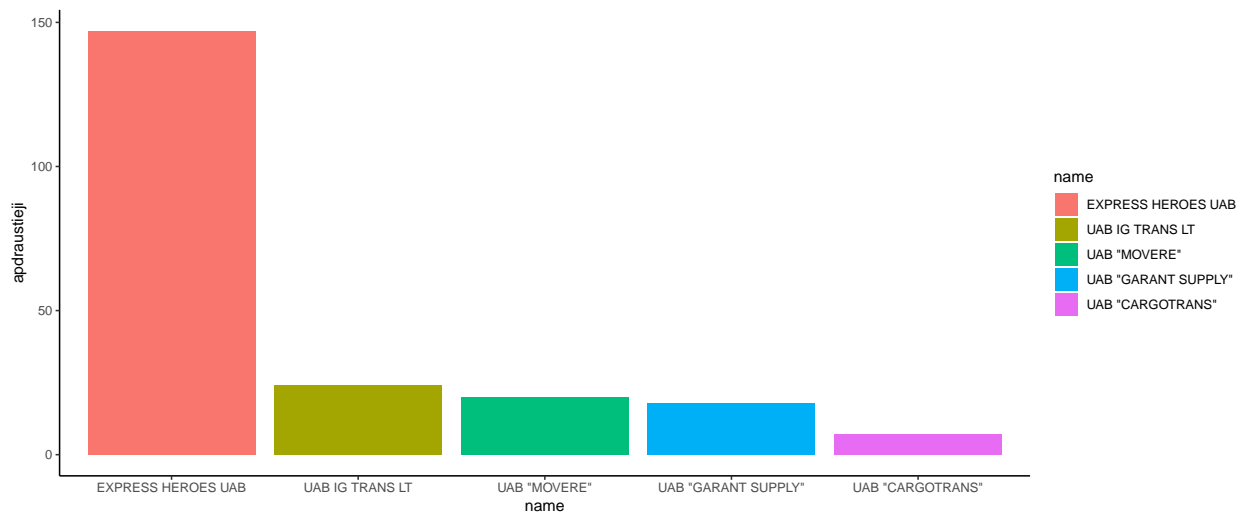
Pavyzdys:



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.

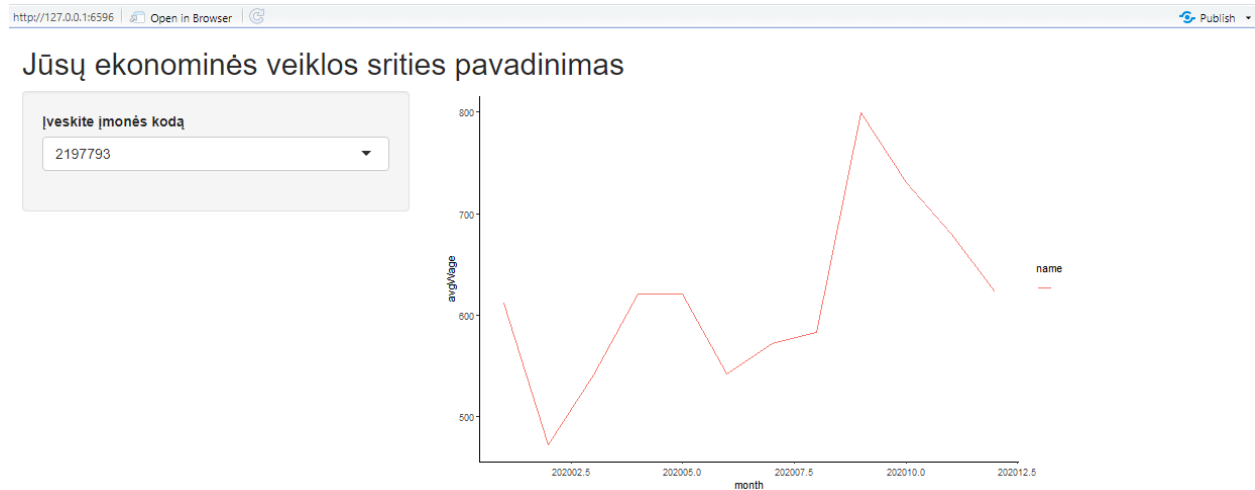


Figure 1: R Shiny aplikacijos pavyzdys