Programists Poll

Penza Street Analitics

23 11 2019

# Опрос программистов от PenzaStreetAnalitics

## Данные

poll <- read.csv("article\_data.csv", encoding = "UTF-8")  
poll["X"] <- NULL  
head(poll)

## gender languages\_number back\_front\_end  
## 1 Кафельный 2 Бек-энд  
## 2 Мужской 1 Бек-энд  
## 3 Мужской 1 Бек-энд  
## 4 Мужской 3 Фул-стек  
## 5 Мужской 2 Фул-стек  
## 6 Мужской 6 Бек-энд

Группой **PenzaStreetAnalitics** была собрана информация об 111 программистах следующего содержания:

1. Пол испытуемого (*gender*)
   * Примечание: в данных имеются записи с “кафельным полом”. Это означает, что программист не счёл важным уточнять свой пол. Для большего охвата, будем учитывать и эту группу.
2. Количество языков программирования (*languages\_number*)
3. А также, какую отрасль web-разработки он бы предпочёл: серверную (*бэк-энд*), клиентскую (*фронт-энд*) или всё сразу (*фул-стек*) (*back\_front\_end*)

## Гипотеза

Влияет ли пол программиста на его выбор отрасли web’а? А также влияет ли выбор отрасли сетевой разработки на количество языков программирования, которые знает программист?

## Исследование

В исследовании использовался язык **R** версии 3.6.1 Также используется библиотека ggplot2

library(ggplot2)

### Зависимость пола и отрасли web’а

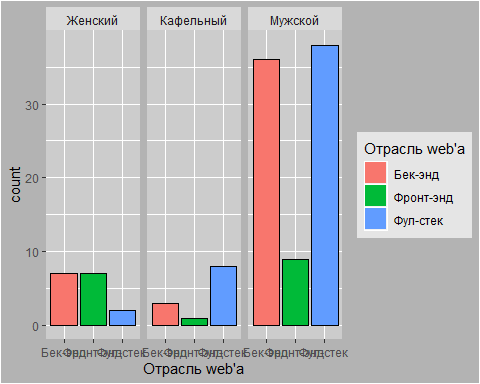
Для проверки гипотезы используем точный тест Фишера:

hyp\_1 <- fisher.test(table(poll$gender, poll$back\_front\_end))  
hyp\_1

##   
## Fisher's Exact Test for Count Data  
##   
## data: table(poll$gender, poll$back\_front\_end)  
## p-value = 0.00486  
## alternative hypothesis: two.sided

p-уровень значимости меньше 0.05 (0.0048598), значит, есть хотя бы одно значимое различие. Визуализируем взаимосвязь:

ggplot(poll, aes(back\_front\_end, fill = back\_front\_end, color = I("black")))+  
 geom\_histogram(stat = "count")+  
 facet\_wrap(~gender)+  
 scale\_fill\_discrete(name = "Отрасль web'а")+  
 xlab("Отрасль web'а")+  
 theme(panel.background = element\_rect(fill = "grey80"),   
 plot.background = element\_rect(fill = "grey70"),  
 legend.background = element\_rect(fill = "grey90"))

 Можно сделать следующие выводы:

1. Частота встречаемости бек-энда и фул-стека у мужчин примерно равна, но фронт-энд встречается реже.
2. Люди с кафельным полом, больше специализирутся на фул-стек разработке.
3. Женщины выбирают фронт-энд чаще, чем остальные. Однако, с такой же частотой он выбирают бек-энд, но редко выбирают фул-стек.

### Зависимость отрасли web’а и количества языков программирования

Проверим нормальность распредения количества языков, с помощью теста Шапиро-Вилка:

norm\_hyp <- shapiro.test(poll$languages\_number)  
norm\_hyp

##   
## Shapiro-Wilk normality test  
##   
## data: poll$languages\_number  
## W = 0.87702, p-value = 4.021e-08

p-уровень значимости меньше 0.05 (4.02145610^{-8}), значит, распределение отличается от нормального. Для проверки гипотезы используем тест Краскела-Уоллеса:

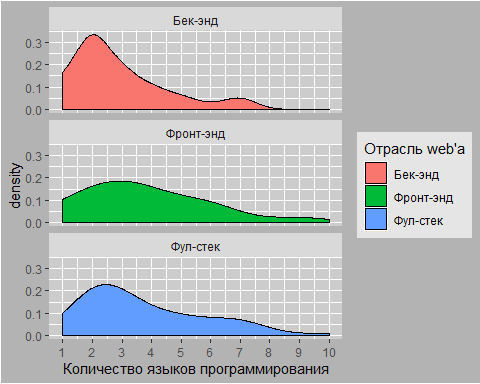
hyp\_2 <- kruskal.test(poll$languages\_number, poll$back\_front\_end)  
hyp\_2

##   
## Kruskal-Wallis rank sum test  
##   
## data: poll$languages\_number and poll$back\_front\_end  
## Kruskal-Wallis chi-squared = 6.0245, df = 2, p-value = 0.04918

p-уровень значимости меньше 0.05 (0.0491798), значит, есть хотя бы одно значимое различие.

Визуализируем взаимосвязь:

ggplot(poll, aes(languages\_number, fill = back\_front\_end, color = I("black")))+  
 geom\_density()+  
 facet\_wrap(~back\_front\_end, nrow = 3)+  
 scale\_x\_continuous(name = "Количество языков программирования",  
 breaks = c(1:10))+  
 scale\_fill\_discrete(name = "Отрасль web'а")+  
 theme(panel.background = element\_rect(fill = "grey80"),   
 plot.background = element\_rect(fill = "grey70"),  
 legend.background = element\_rect(fill = "grey90"))

 Можно сделать следующие выводы:

1. У всех трёх отраслей наибольшее число программистов знают 2-3 языка.
2. Для фронт-энда прослеживается большая частота встречаемости программистов, знающих >7 языков по сравнению с бек-эндом. Такое же различие заметно для фул-стека.

## Вывод

Программисты - странные люди. Как оказалось, в их обществе тоже существуют свои закономерности и правила. Их исследованию можно уделить множество статей. Но как-нибудь в другой раз.