

## ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

(приклад виконання)

У файлі `OpenData2018.csv` міститься інформація про результати незалежного зовнішнього тестування 2018 року (Джерело даних: <https://www.kaggle.com/ipopovych/ukrainian-independent-testing-results-2018/downloads/ukrainian-independent-testing-results-2018.zip/2>).

Перед початком роботи підключіть бібліотеки *dplyr*, *ggplot2*, *readr*.

Завантажимо файл із папки ZNO

```
OpenData2018 <- read_delim("ZNO/OpenData2018.csv", delim = ";", escape_double = FALSE, trim_ws = TRUE)
```

Розміщення файлу на диску виділене курсивом замініть на власну адресу.

Для аналізу інформації про результати незалежного зовнішнього тестування 2018 року була обрана Харківська область. Запишемо дані про вступників Харківської області у нову змінну **znoKharkiv**.

```
znoKharkiv <- OpenData2018[OpenData2018$REGNAME == 'Харківська область', ]
```

Збережемо дані у окремому файлі **znoKharkiv.csv**.

```
write.csv(znoKharkiv, "znoKharkiv.csv")
```

Щоб прочитати файл з даними Харківської області виконайте прочитання з файлу

```
znoKharkiv <- read.csv('znoKharkiv.csv')
```

Перегляньте дані

```
view(znoKharkiv)
```

X	OUTID	Birth	SEXTYPENAME	REGNAME
1	6504a7f3-7dfd-4a64-a2d9-4d49f5021303	2000	жіноча	Харківськ
2	0328e5df-4915-436c-bd88-d23ebf958073	2000	чоловіча	Харківськ
3	14ee428b-9948-4c43-9c81-ec2cbba305cc	1999	жіноча	Харківськ
4	319a5e74-8b33-411b-853b-07f67585e6ca	2001	чоловіча	Харківськ
5	adf61abf-6e4c-4323-b7cf-f696f0ddbc4f	1999	жіноча	Харківськ
6	2eb51dbf-5b73-4f79-96a8-930de70bbffa	2000	чоловіча	Харківськ
7	91e35584-ad21-41a3-af5e-0c84a9bd0fb2	2000	жіноча	Харківськ
8	5eebdd73-eda1-4c0e-963e-20ec4018e817	2001	жіноча	Харківськ
9	84dcf1bd-2676-4366-b505-92557795ef9f	2001	чоловіча	Харківськ
10	7ad7fc38-bea7-46f8-8fd2-b6c10dac89fd	2001	жіноча	Харківськ
11	faf136df-c249-4126-a8d9-f5cc5b9530a6	2000	жіноча	Харківськ
12	22b69393-3373-4428-ab17-950614acd2cd	2000	чоловіча	Харківськ
13	43857756-7974-4f47-87b4-c12bdad1e9b4	1998	чоловіча	Харківськ
14	eca160de-f816-4150-8351-ceff4edealb6	2001	чоловіча	Харківськ
15	fe48112c-041e-4618-82a1-76cb486656ca	2000	чоловіча	Харківськ

Підрахунок кількості рядків

```
nrow(znokharkiv)
```

```
[1] 21251
```

Можна зробити висновок, що у складанні ЗНО по Харківській області взяло участь 21251 осіб (не враховуючи учасників “null” про яких відсутня будь яка інформація).

Проведемо аналіз учасників за профілем навчання.

Запишемо дані про профіль навчання у змінну *status*

```
status <- table(znokharkiv$ClassProfileNAME)
status
```

```

null
9636
Біолого-хімічний
538
Біотехнологічний
153
Військово-спортивний
149
Географічний
15
Екологічний
227
Економічний
641
Іноземної філології
1488
Інформаційно-технологічний
504
Інший(багатопротильність)
111
Історичний
```

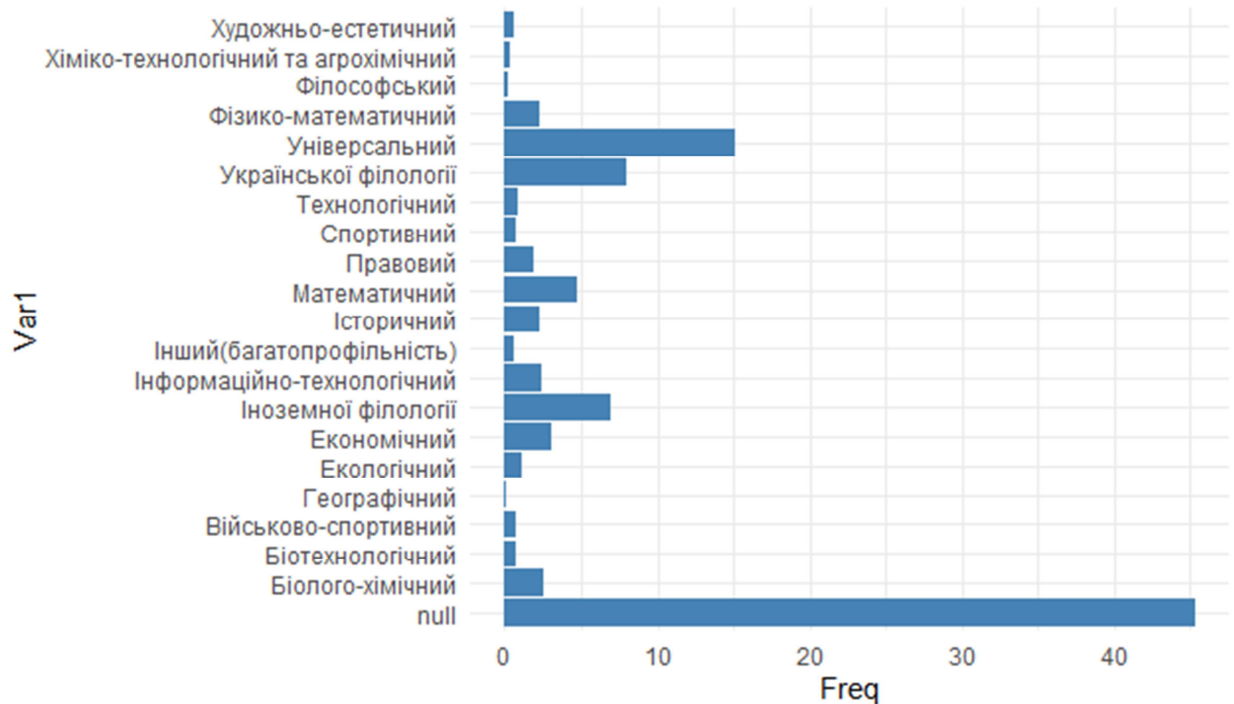
	474
Математичний	1000
Правовий	388
Спортивний	147
Технологічний	178
Української філології	1687
Універсальний	3206
Фізико-математичний	471
Філософський	40
Хіміко-технологічний та агрохімічний	63
Художньо-естетичний	135

Розрахуємо відсоткові співвідношення профілів навчання

```
perc <- status/sum(status) * 100
perc
```

	perc
Біолого-хімічний	45.34374853
Біотехнологічний	2.53164557
Військово-спортивний	0.71996612
Географічний	0.70114348
Екологічний	0.07058491
Економічний	1.06818503
Іноземної філології	3.01632864
Інформаційно-технологічний	7.00202343
Інший (багато профільність)	2.37165310
Історичний	0.52232836
Математичний	2.23048327
Правовий	4.70566091
Спортивний	1.82579643
Технологічний	0.69173215
Української філології	0.83760764
Універсальний	7.93844996
Фізико-математичний	15.08634888
Філософський	2.21636629
Хіміко-технологічний та агрохімічний	0.18822644
Художньо-естетичний	0.29645664
	0.63526422

```
p<-data.frame(perc)
ggplot(data=p, aes(x=Var1, y=Freq)) +geom_bar(stat="identity", fill="steelblue")
+ theme_minimal()+coord_flip()
```



**В результаті аналізу можна зробити висновок, що найбільшими (за кількістю учасників) є універсальний профіль, іноземної філології, української філології та математичний. У 9639 осіб інформація про тип профілю відсутня.**

### **Розрахуємо статистичні характеристики балів учасників ЗНО з української мови по Харківській області**

Позбавимось записів з відсутніми оцінками з української мови

```
kharkivua<-znoKharkiv[znoKharkiv$UkrBall100!='null',]
```

Розрахуємо середнє значення:

```
mean(as.numeric(kharkivua$UkrBall100))
[1] 128.4766
```

```
median(as.numeric(kharkivua$UkrBall100))
[1] 138
```

Розрахуємо значення середньоквадратичного відхилення

```
sd(as.numeric(kharkivua$UkrBall100))
[1] 52.89775
```

Розрахуємо кuartиль:

```
quantile(as.numeric(kharkivua$UkrBall100))
```

0%	25%	50%	75%	100%
0.0	114.0	138.0	165.0	199.5

**Можна зробити висновок, що 75% учасників склали ЗНО з Української мови на 165.0 тестові бали з 200 і лише 1 людина склала на 199,5 що є максимальним результатом по області.**