

Лабораторна робота №4. Варіант 15

Виконав: студент групи ІП-13 Ал Хадам Мурат Резгович

Візуалізація даних за допомогою matplotlib та Seaborn

Мета роботи: Ознайомитись з основними діаграмами та графіками, що використовуються при аналізі даних. Навчитись будувати їх за допомогою бібліотек matplotlib та Seaborn.

Завдання до лабораторної роботи:

- Побудувати стовпчикові діаграми, на яких відобразити: а) кількість клієнтів з підключеною голосовою поштою та без; б) мінімальну кількість хвилин в вечірній час для клієнтів з підключеною голосовою поштою і без; в) середню оплату міжнародних дзвінків для клієнтів з підключеним роумінгом та без і з врахуванням лояльності.
- Побудувати гістограму хвилин в нічний час, загальну і в залежності від підключеного роумінгу.
- Побудувати діаграму розмаху дзвінків в денний час (загальну і в залежності від підключеної голосової пошти), визначити чи присутні викиди.
- За допомогою діаграм розсіювання зробити висновки щодо залежності між: а) міжнародними хвилинами і дзвінками; б) міжнародними хвилинами і оплатою. Порахувати коефіцієнт кореляції за допомогою відповідних функцій.

```
In [1]:
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('telecom.csv')
df

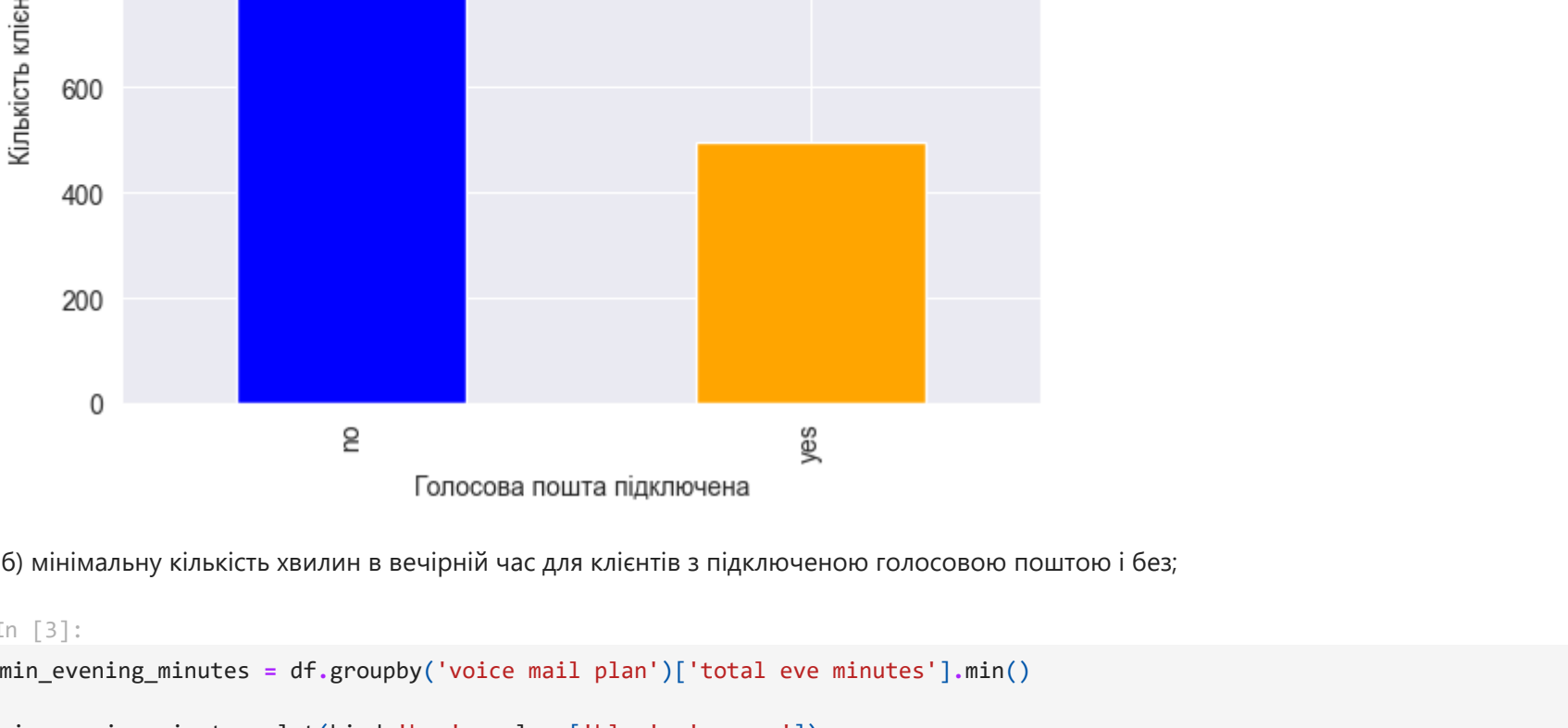
Out[1]:
```

	state	account length	area code	phone number	international plan	voice mail plan	number vmail messages	total day minutes	total day calls	total day charge	...	total eve calls	total eve charge	total night minutes	total night calls
0	KS	128	415	382-4657	no	yes	25	265.1	110	45.07	...	99	16.78	244.7	91
1	OH	107	415	371-7191	no	yes	26	161.6	123	27.47	...	103	16.62	254.4	103
2	NJ	137	415	358-1921	no	no	0	243.4	114	41.38	...	110	10.30	162.6	104
3	OH	84	408	375-9999	yes	no	0	299.4	71	50.90	...	88	5.26	196.9	89
4	OK	75	415	330-6626	yes	no	0	166.7	113	28.34	...	122	12.61	186.9	121
...
1730	NM	161	408	397-8011	no	no	0	107.5	121	18.28	...	46	21.79	247.2	131
1731	TX	70	415	341-8719	no	no	0	232.8	95	39.58	...	111	25.79	255.6	104
1732	HI	97	415	408-1242	no	yes	43	121.1	105	20.59	...	115	22.12	222.4	100
1733	WA	130	510	406-7726	no	no	0	124.3	70	21.13	...	99	23.01	239.5	83
1734	WI	92	415	351-2773	no	no	0	157.7	101	26.81	...	100	25.38	216.9	99

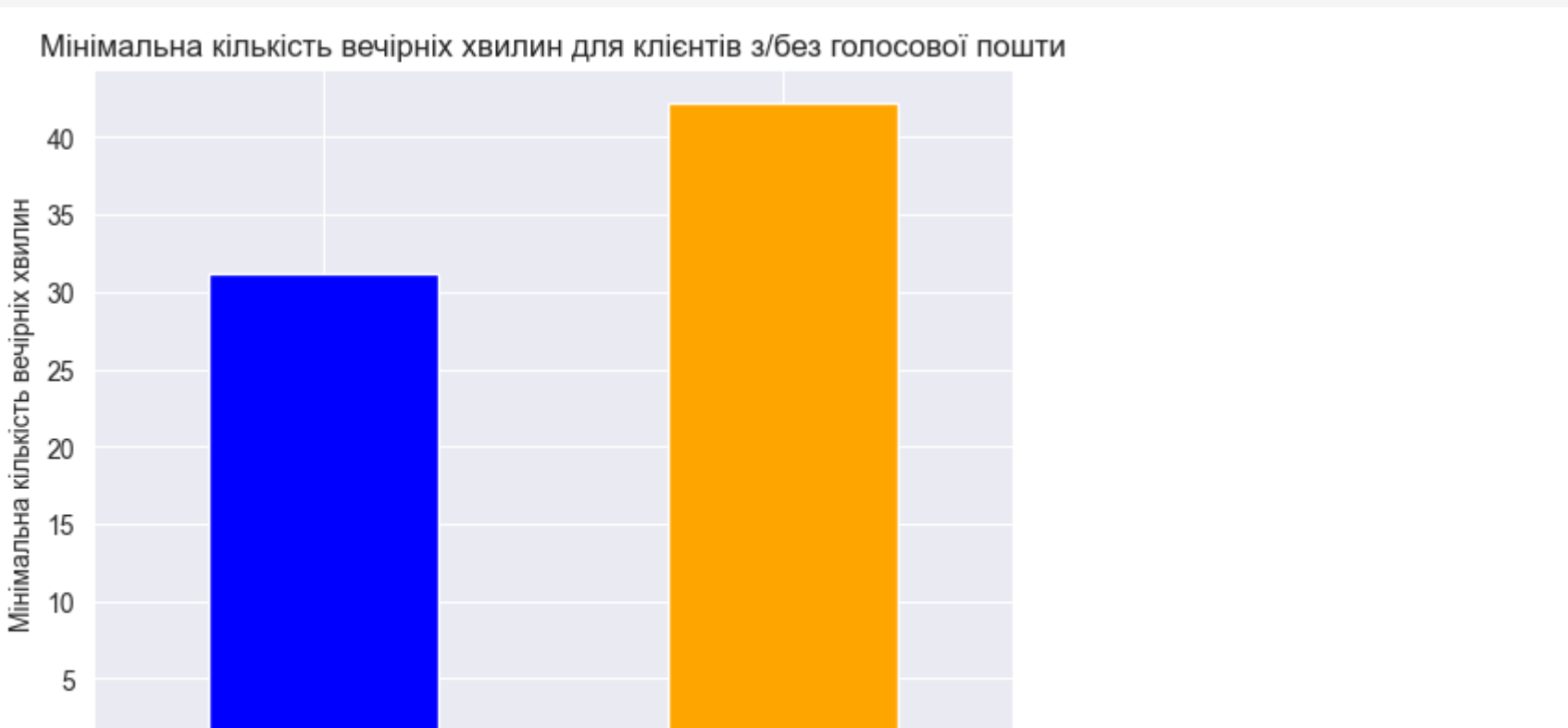
1735 rows × 16 columns

1. Побудувати стовпчикові діаграми, на яких відобразити:

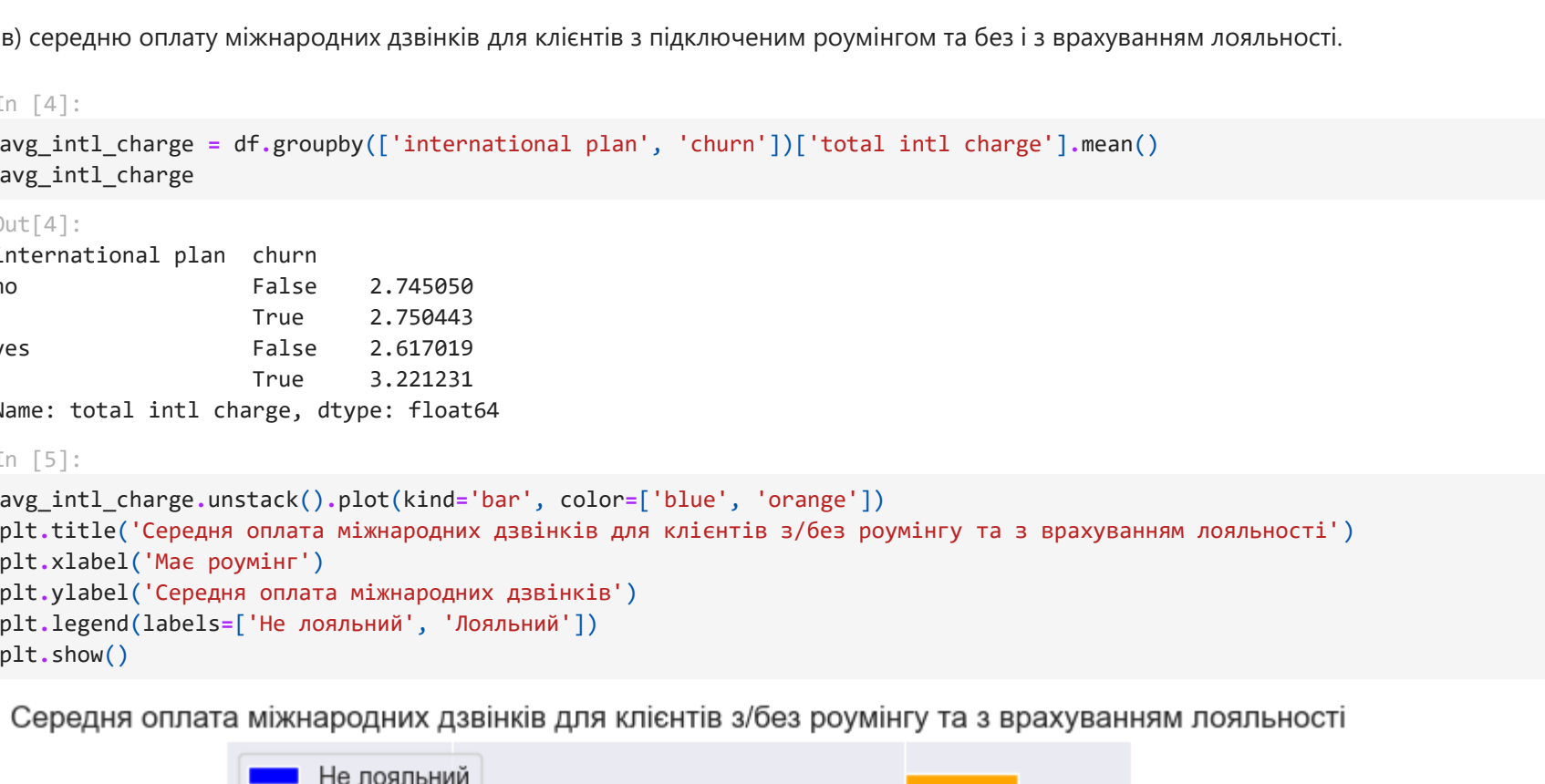
а) кількість клієнтів з підключеною голосовою поштою та без;



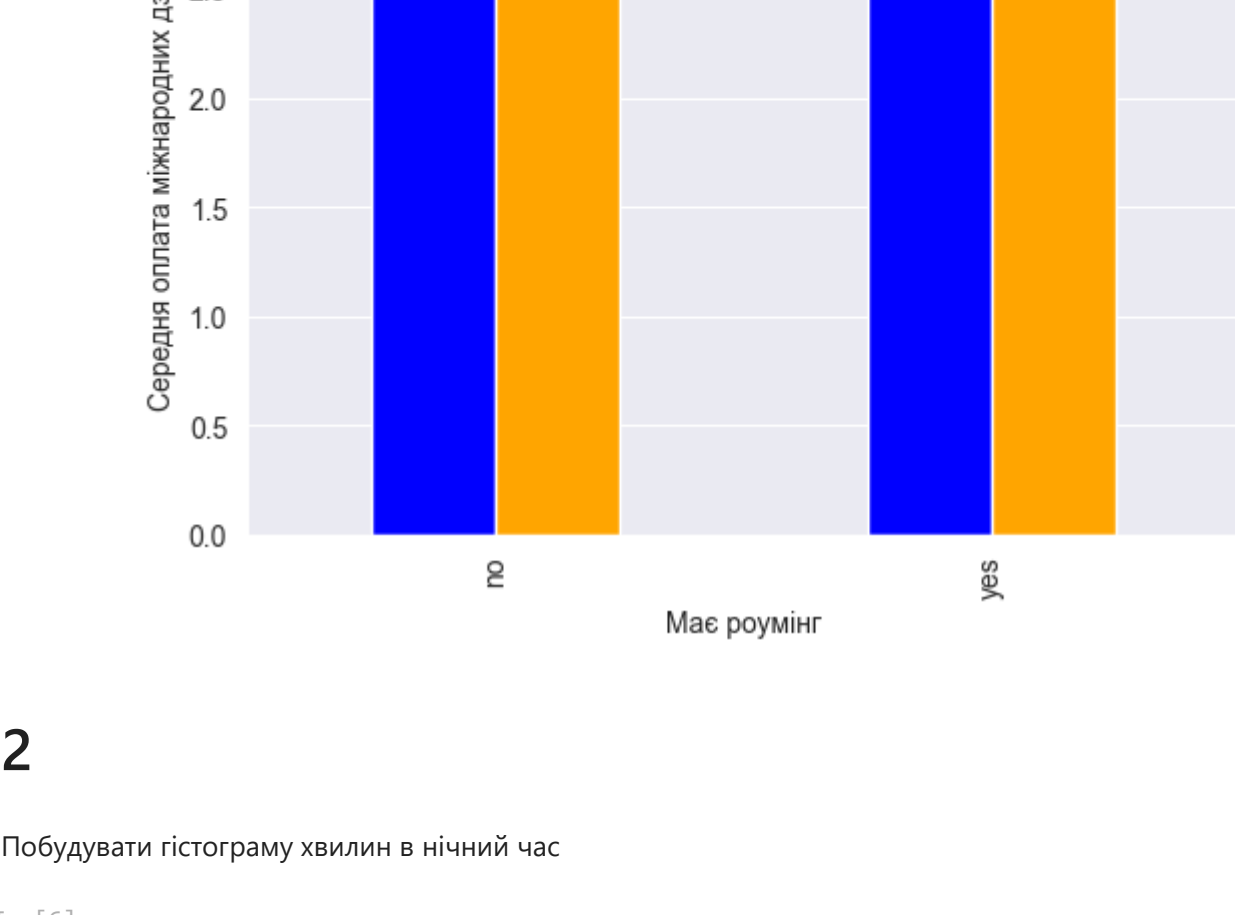
б) мінімальну кількість хвилин в вечірній час для клієнтів з підключеною голосовою поштою і без;



в) середню оплату міжнародних дзвінків для клієнтів з підключеним роумінгом та без і з врахуванням лояльності.

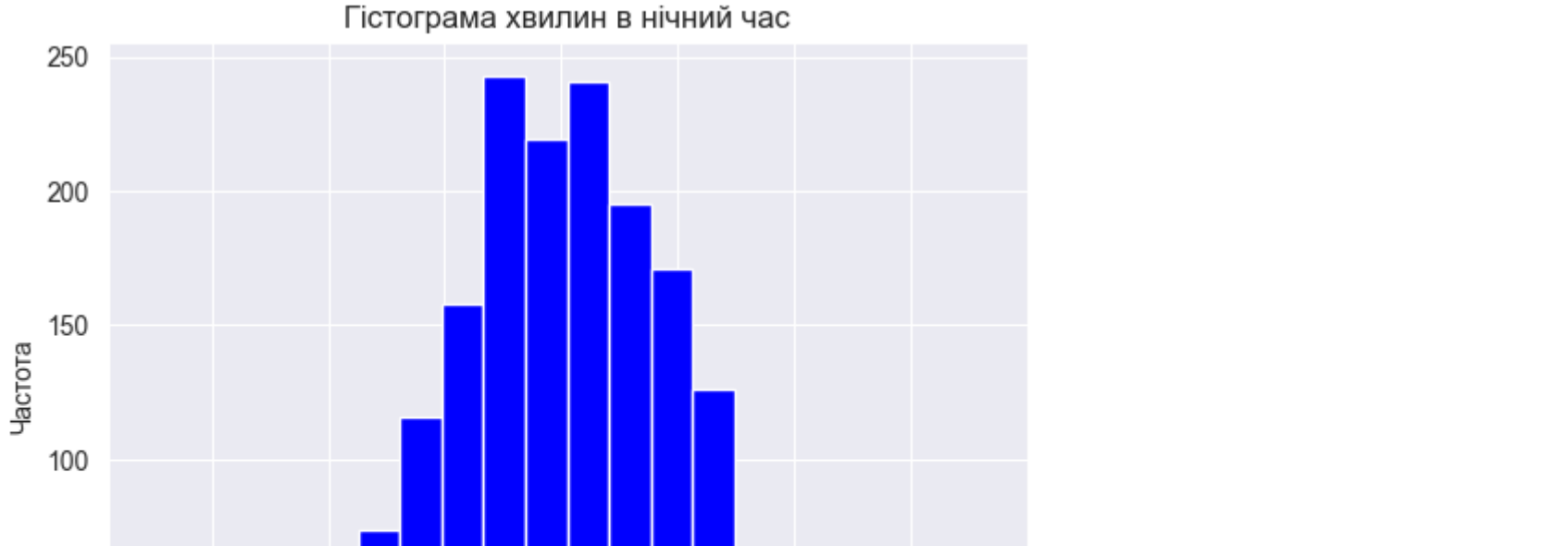


Середня оплата міжнародних дзвінків для клієнтів з/без роумінгу та з врахуванням лояльності

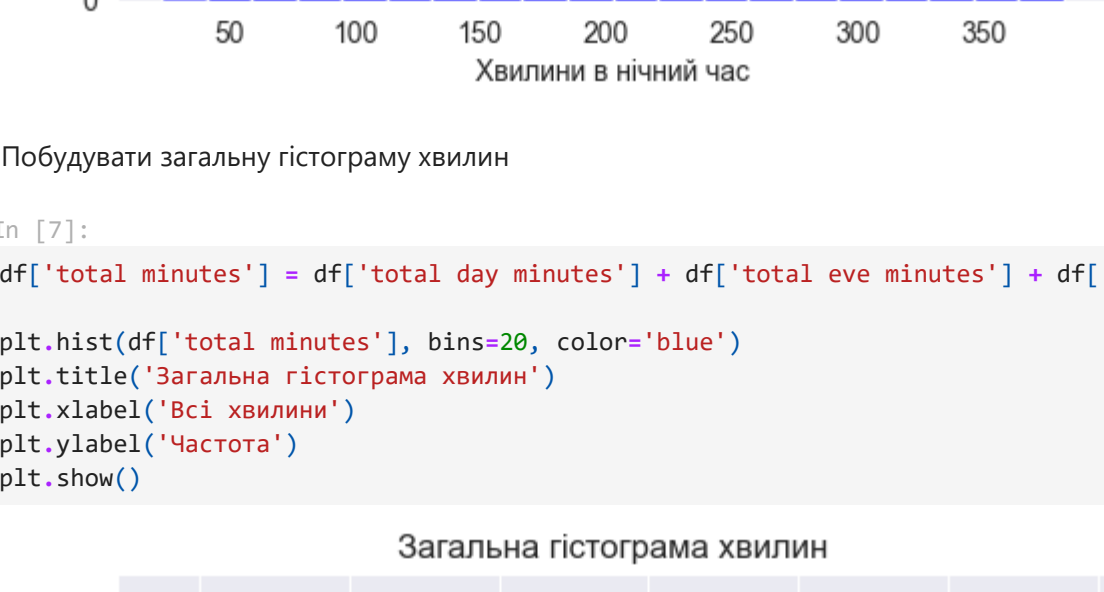


2

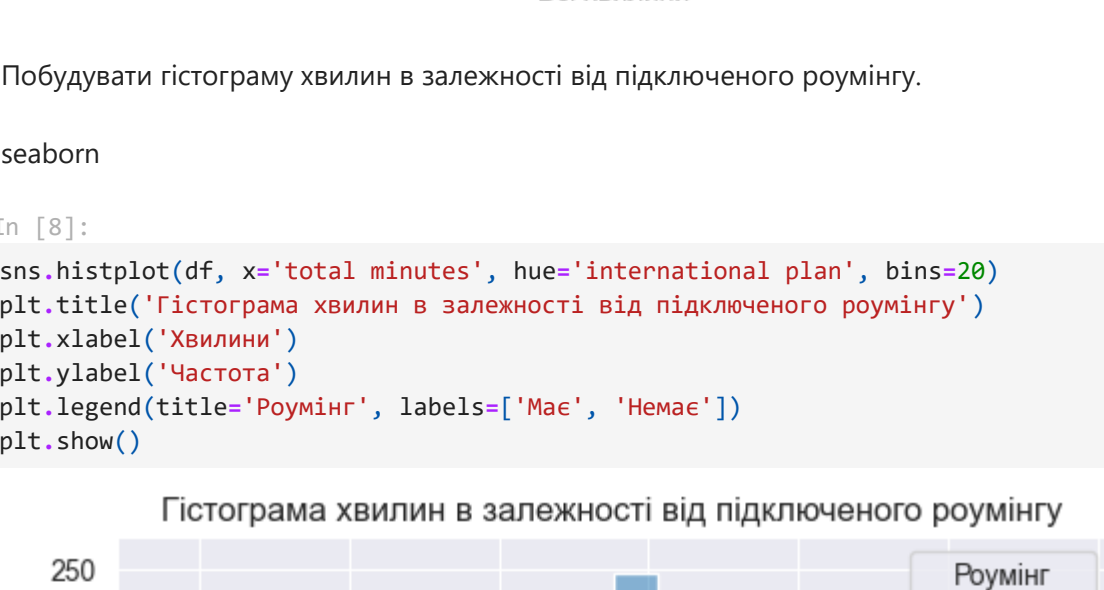
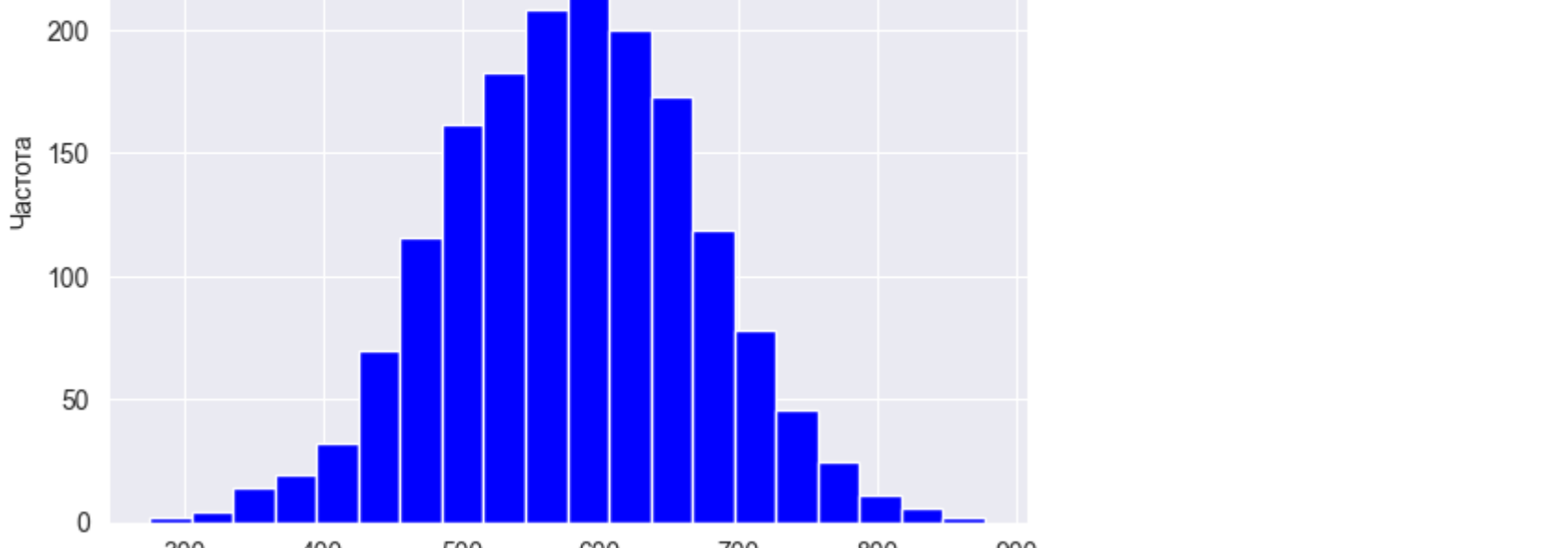
Побудувати гістограму хвилин в нічний час



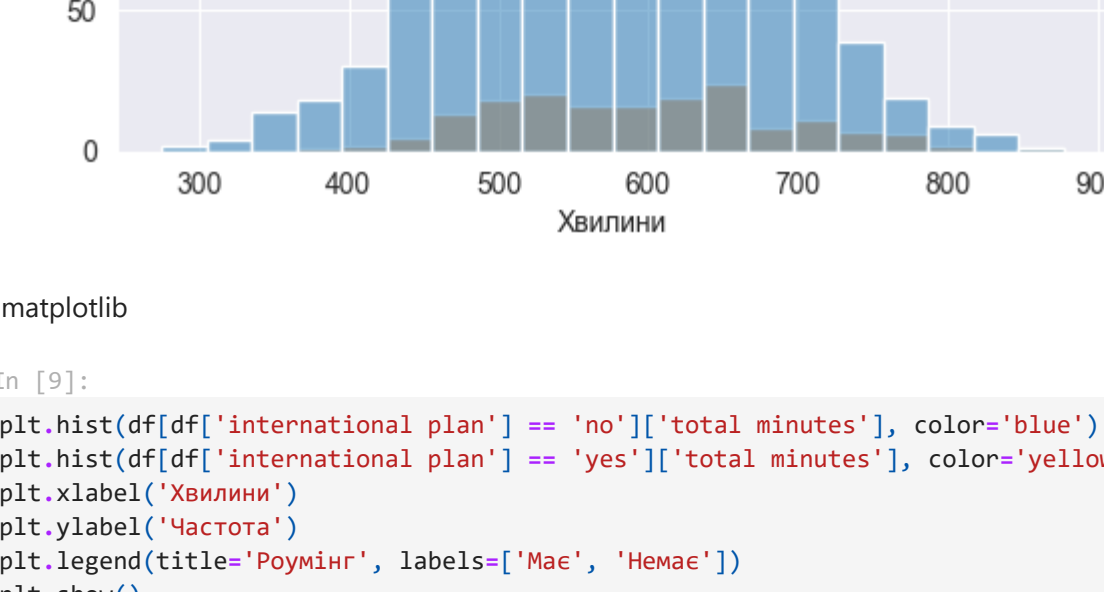
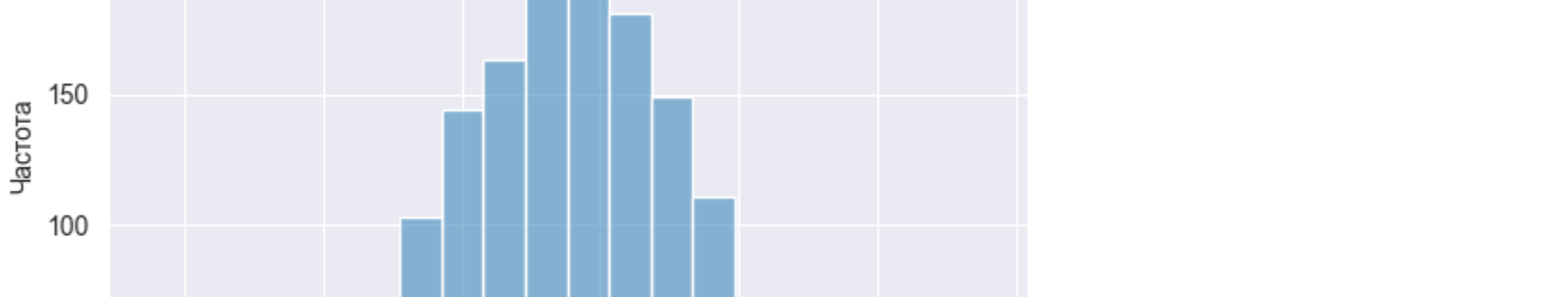
Гістограма хвилин в нічний час



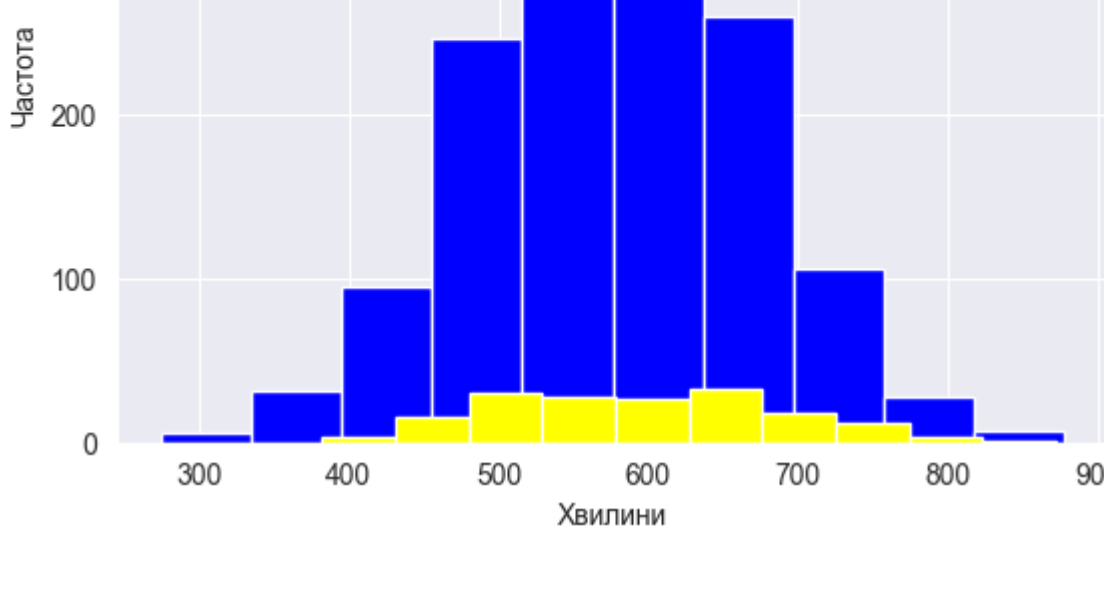
Побудувати загальну гістограму хвилин



Побудувати гістограму хвилин в залежності від підключеного роумінгу.

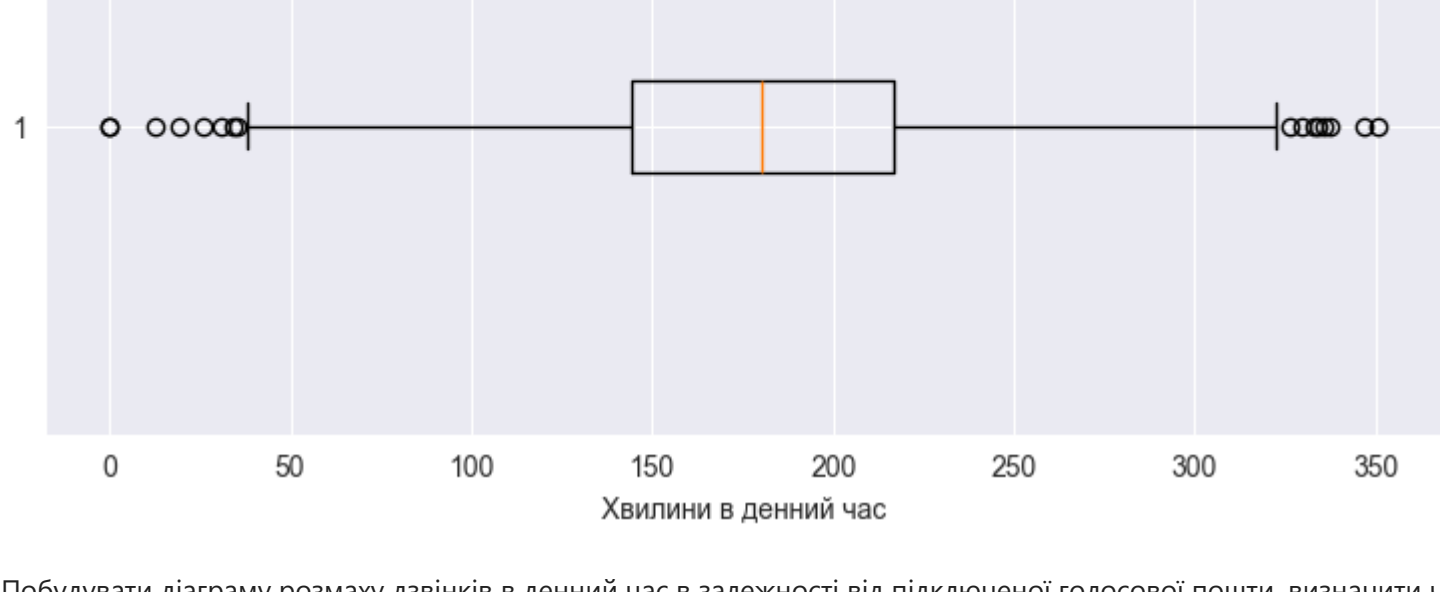
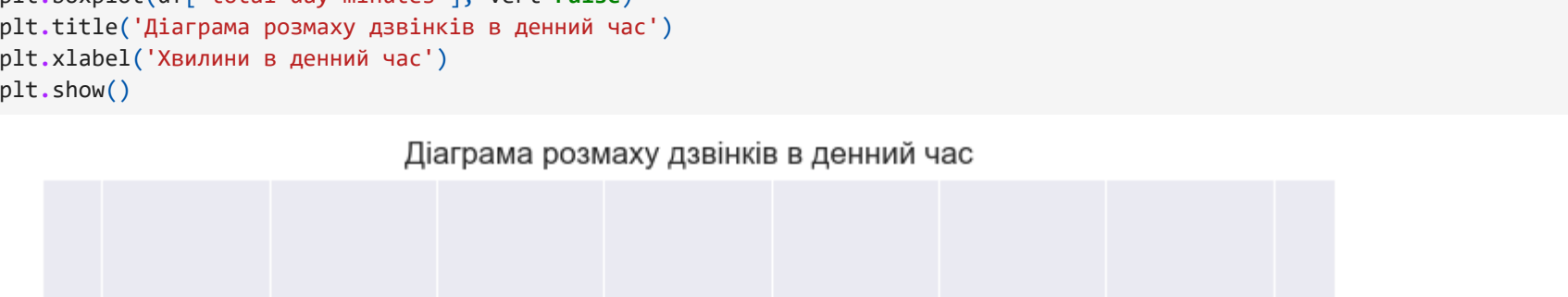


Гістограма хвилин в залежності від підключеного роумінгу

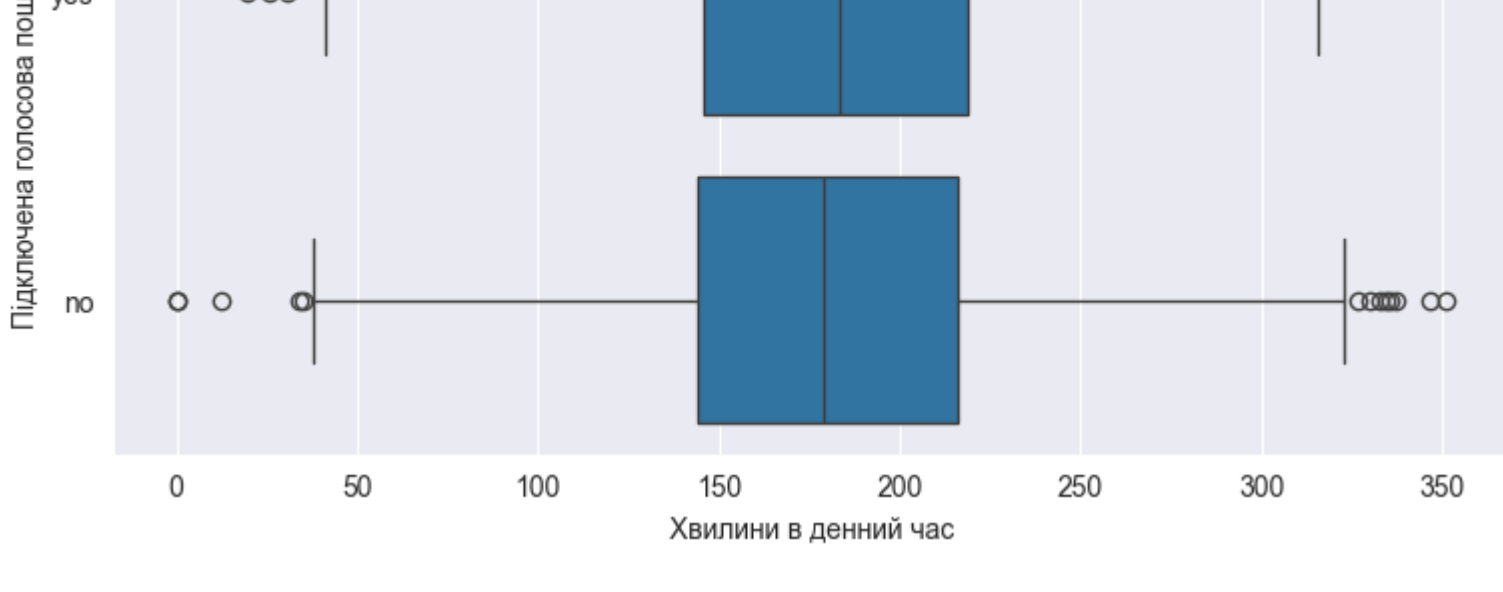
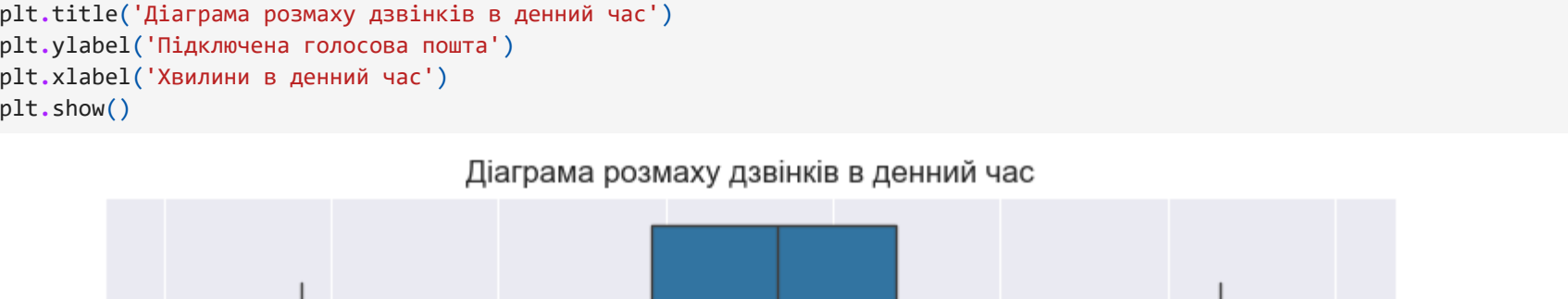


3

Побудувати діаграму розмаху дзвінків в денний час загальну, визначити чи присутні викиди.



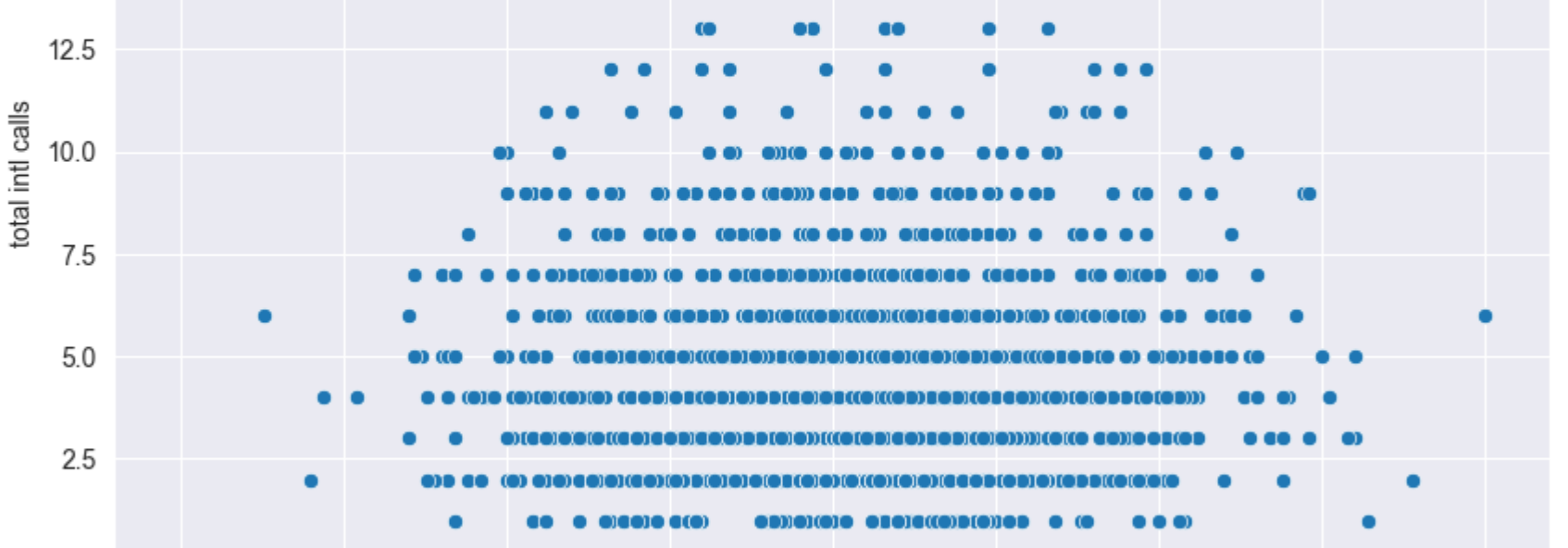
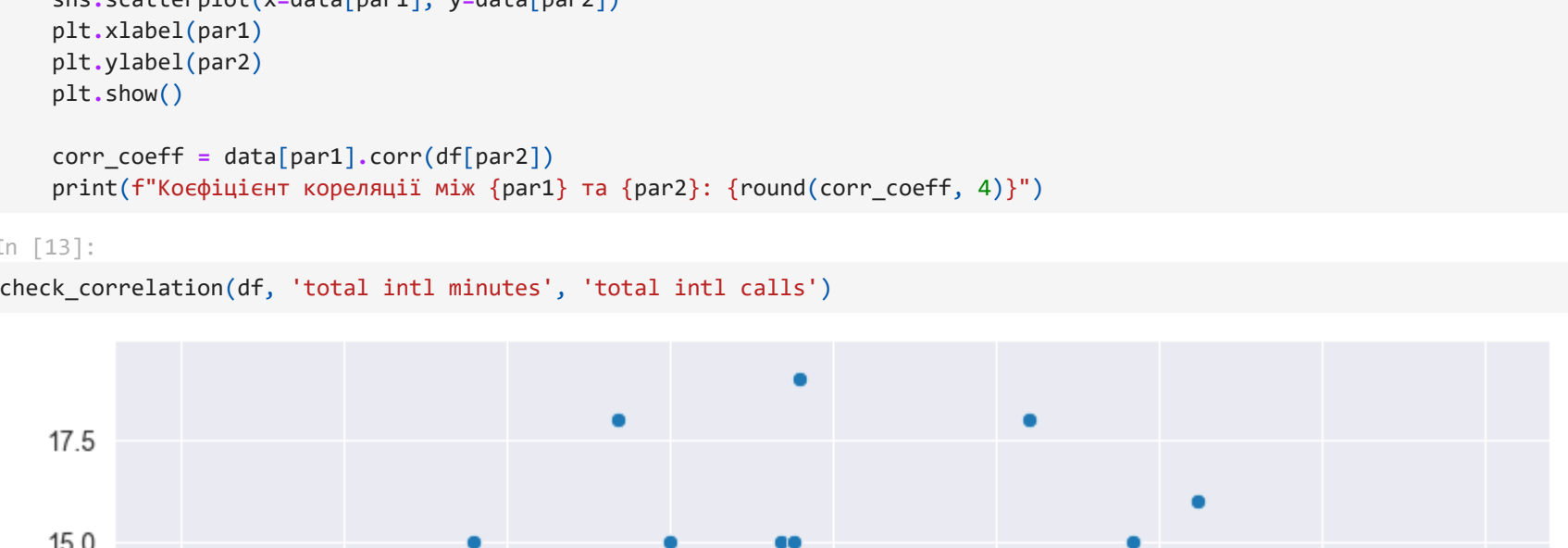
Побудувати діаграму розмаху дзвінків в денний час в залежності від підключеної голосової пошти, визначити чи присутні викиди.



Викиди позначені на діаграмі розкиду поза вусами boxplot.

4

За допомогою діаграм розсіювання зробити висновки щодо залежності між: а) міжнародними хвилинами і дзвінками;

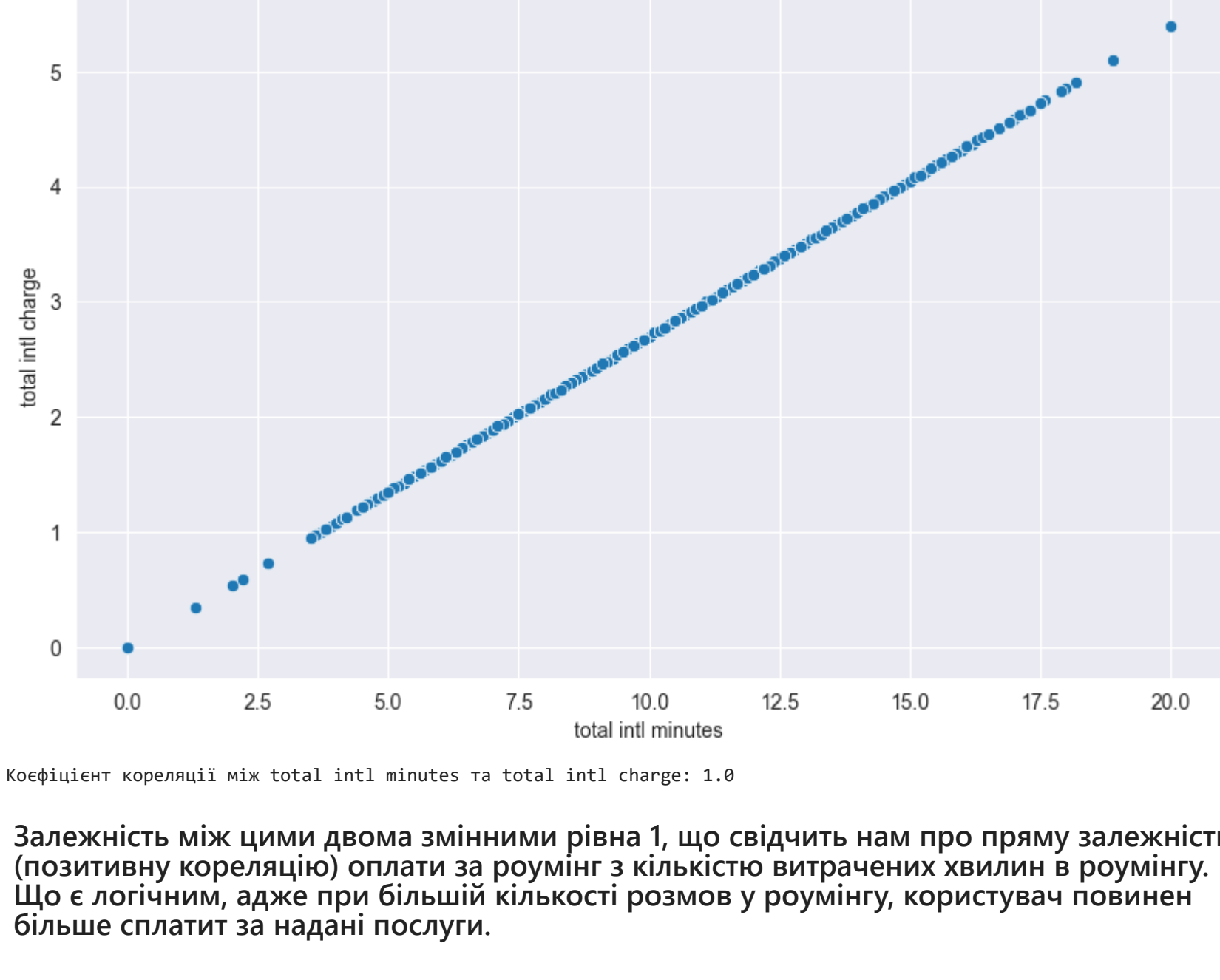


Коефіцієнт кореляції між total intl minutes та total intl calls: 0.0487

Залежність між цими двома змінними майже відсутня, що можна побачити, і на діаграмі розсіювання, і за значенням коефіцієнту кореляції.

б) міжнародними хвилинами і оплатою.





Коефіцієнт кореляції між total intl minutes та total intl charge: 1.0

Залежність між цими двома змінними рівна 1, що свідчить нам про пряму залежність (позитивну кореляцію) оплати за роумінг з кількістю витрачених хвилин в роумінгу. Що є логічним, адже при більшій кількості розмов у роумінгу, користувач повинен більше сплатити за надані послуги.