МІНІСТЕРСТВО ОСВІТИ І НАУКИ

ДВНЗ “КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ

імені Вадима Гетьмана”

Кафедра комп'ютерної математики та інформаційної безпеки

**Звіт**

**до лабораторної роботи № 12**

“Візуалізація даних: pandas і matplotlib”

**Виконала:**

Кунєва К.Р., ІА-301

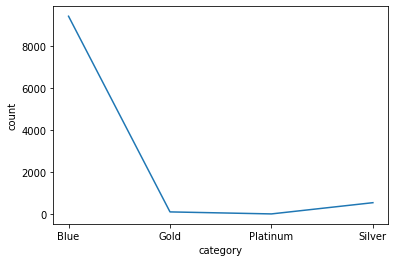
**Перевірила:**

доц. Корзаченко О.В.

Київ — 2022

**Аналіз даних на основі побудови графіків з використанням matplotlib.**

1. Лінійна діаграма



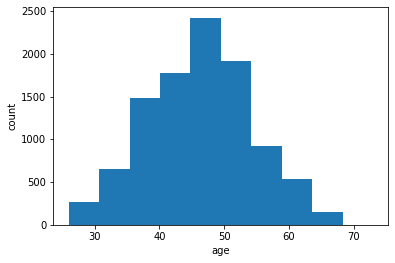
Найбільша кількість клієнтів має картки категорії «Blue», інші категорії використовуються набагато рідше і між ними, як видно з графіку, невеликі розриви. А саме: картки категорії «Platinum» має найменша кількість клієнтів, наступними є категорії «Gold» і «Silver» відповідно.

1. Кругова діаграма



На круговій діаграмі відображені співвідношення вікових категорій клієнтів, починаючи від 25 років із кроком 10. Найбільша кількість клієнтів знаходиться в категорії від 45 до 50, потім – від 35 до 45, від 55 до 65 і від 25 до 35.

1. Розподіл вв



Гістограма показує дані, згруповані в 10 стовпчиків від 20 до 70 (у роках) і кожен стовпчик має ширину 5. Випадкова величина «Вік» не має нормального розподілу.

Гістограма досягає піку зліва нижче 50 років, хвости від нього простягаються в обидві сторони, і вправо, і вліво, але з різними кількостями. Це означає, що клієнтів молодше 45 років більше, ніж тих, хто старше 50, але розрив не великий.

1. Кореляція між даними

Кореляція даних відносно обраної вв («віку»). Значення кореляції віку з самим собою 1, це означає лінійну залежність. Чим ближе значення до 0, тем менша кореляція. Знак (“+” або “-”) означає зростання або спадання значень відносно значень випадкової величини.

