**Міністерство освіти і науки України**

**Національний технічний університет України «КПІ» імені Ігоря Сікорського**

**Кафедра обчислювальної техніки ФІОТ**

**ЗВІТ**

**з лабораторної роботи №5**

**з навчальної дисципліни «Вступ до технології Data Science»**

**Тема:**

**ДОСЛІДЖЕННЯ ТЕХНОЛОГІЙ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ**

**Виконав:**

Студент X курсу кафедри ОТ ФІОТ,

Навчальної групи ІА-73

Петренко П.П.

**Перевірив:**

Професор кафедри ОТ ФІОТ

Писарчук О.О.

**Київ 2023**

**І. Мета:**

виявити дослідити та узагальнити особливості інтелектуального аналізу даних з використанням спеціалізованих пакетів мови програмування Python.

**ІІ. Завдання:**

| 6, 21 | Розробити програмний скрипт, що реалізує оцінювання ефективності впровадження нового товару на ринку продукції. Вихідні дані містять 12 критеріїв, з яких 7 – максимізованих, решта – мінімізованих. Кількість аналогічних товарів – 8. Вхідні дані занесені у файл. Формат файлу, тип товару, зміст показників та критеріїв ефективності та їх значення обрати самостійно.  ***З використанням OLAP технологій*** провести інтелектуальний аналіз результатів розв’язку багатокритеріальної задачі, довести їх адекватність та сформувати аналітичний звіт за результатами досліджень. |
| --- | --- |

Використання OLAP для цієї лабораторної роботи не має сенсу, оскільки цей пакет застарілий та має набагато кращу, підтримувану, актуальну та використовувану альтернативу - pandas та numpy.

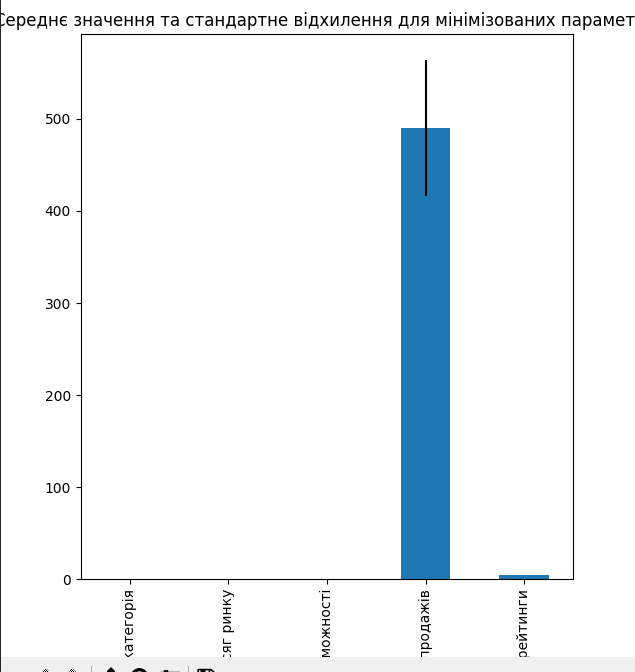


Рисунок 1.

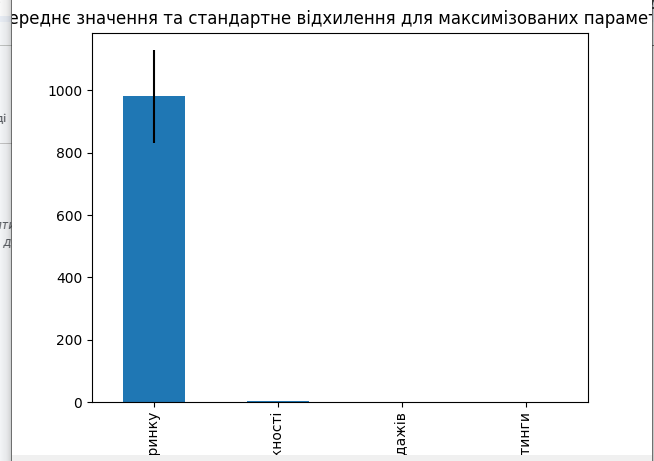


Рисунок 2.

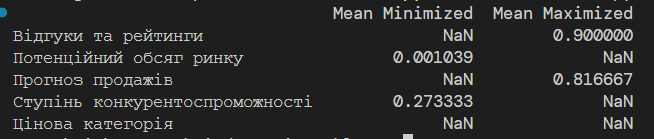


Рисунок 3. аналіз виходу на ринок товарів вцілому

**ІІІ. Довести адекватність сформованих моделей та працездатність розробленого скріпта.**

**Розроблений код повинен бути раціональним та відповідати вимогам до чистого коду.**

**import pandas as pd**

**import numpy as np**

**import matplotlib.pyplot as plt**

**def create\_dataset():**

**# Згенеруємо приклади для датасету**

**np.random.seed(42)**

**data = {**

**'Назва товару': ['Товар1', 'Товар2', 'Товар3', 'Товар4', 'Товар5'],**

**'Категорія товару': ['Електроніка', 'Одяг', 'Харчові продукти', 'Електроніка', 'Одяг'],**

**'Цінова категорія': ['Середня', 'Висока', 'Низька', 'Середня', 'Висока'],**

**'Цінова стратегія': ['Знижки', 'Пакетні пропозиції', 'Знижки', 'Знижки', 'Пакетні пропозиції'],**

**'Цільова аудиторія': ['Молодь', 'Дорослі', 'Сім\'я', 'Молодь', 'Дорослі'],**

**'Конкуренти': ['Конкурент1', 'Конкурент2', 'Конкурент3', 'Конкурент1', 'Конкурент2'],**

**'Маркетингові стратегії': ['Реклама', 'Соціальні мережі', 'Реклама', 'Реклама', 'Соціальні мережі'],**

**'Потенційний обсяг ринку': [1000, 800, 1200, 900, 1000],**

**'Інноваційні рішення': ['Технологія A', 'Дизайн B', 'Технологія A', 'Технологія C', 'Дизайн B'],**

**'Ступінь конкурентоспроможності': [4, 3, 5, 4, 3],**

**'Прогноз продажів': [500, 400, 600, 450, 500],**

**'Відгуки та рейтинги': [4.5, 4.0, 4.8, 4.2, 4.1]**

**}**

**df = pd.DataFrame(data)**

**df.to\_csv('variant6.csv', index=False)**

**create\_dataset()**

**df = pd.read\_csv('variant6.csv')**

**minimize\_criteria = ['Цінова категорія', 'Потенційний обсяг ринку', 'Ступінь конкурентоспроможності']**

**maximize\_criteria = ['Прогноз продажів', 'Відгуки та рейтинги']**

**# Мінімізуємо параметри**

**df\_minimized = df.copy()**

**df\_minimized['Цінова категорія'] = pd.to\_numeric(df\_minimized['Цінова категорія'], errors='coerce')**

**df\_minimized[minimize\_criteria] = df\_minimized[minimize\_criteria].apply(lambda x: 1 / x)**

**# Максимізуємо параметри**

**df\_maximized = df.copy()**

**df\_maximized[maximize\_criteria] = df\_maximized[maximize\_criteria].apply(lambda x: x / x.max())**

**# Зберігаємо результат у нові файли**

**df\_minimized.to\_csv('minimized\_data.csv', index=False)**

**df\_maximized.to\_csv('maximized\_data.csv', index=False)**

**df\_minimized.describe().transpose()[['mean', 'std']].plot(kind='bar', y='mean', yerr='std', legend=False)**

**plt.title('Середнє значення та стандартне відхилення для мінімізованих параметрів')**

**plt.show()**

**df\_maximized.describe().transpose()[['mean', 'std']].plot(kind='bar', y='mean', yerr='std', legend=False)**

**plt.title('Середнє значення та стандартне відхилення для максимізованих параметрів')**

**plt.show()**

**mean\_minimized = df\_minimized.describe().transpose()['mean']**

**mean\_maximized = df\_maximized.describe().transpose()['mean']**

**analysis\_df = pd.DataFrame({**

**'Mean Minimized': mean\_minimized[minimize\_criteria],**

**'Mean Maximized': mean\_maximized[maximize\_criteria],**

**})**

**print(analysis\_df)**

**III. Висновок:**

виявив дослідив та узагальнив особливості інтелектуального аналізу даних з використанням спеціалізованих пакетів мови програмування Python.

Виконав: студент Петренко П.П.