## Розділ 8. Вибір і застосування методів інтелектуального аналізу даних(ІАД) для управління бізнес-процесами

* 1. **Опис методів ІАД та обґрунтування доцільності їх використання**

Кластерний аналіз представляє собою ефективний метод для групування хлібобулочних виробів на однорідні категорії за виробничими та економічними параметрами. Якщо дані про продукцію представити у вигляді точок у просторі характеристик, то завдання кластеризації полягає у виявленні природних "концентрацій точок".

Іншими словами, кластерний аналіз — це аналітична процедура, що здійснює багатопараметричну класифікацію даних про хлібобулочну продукцію, які містять інформацію про вибірку виробів, тобто систематизує продукти у відносно однорідні категорії. Таким чином вирішується завдання класифікації асортименту з використанням чіткого математичного апарату.

Кластеризація є дескriptивною процедурою, вона не формує статистичних висновків, але дає змогу провести розвідувальний аналіз і вивчити "структуру асортименту".

Мета кластерного аналізу — розподіл видів хлібобулочних виробів на відносно однорідні кластери, виходячи з розглянутого набору змінних, таким чином, щоб в один кластер потрапляли подібні за характеристиками продукти, а в різні кластери — віддалені один від одного за властивостями вироби. Іншими словами, кожному кластеру відповідає концентрація точок у просторі змінних, а різним кластерам — різні концентрації, розділені більшими або меншими інтервалами.

**Методи кластерного аналізу можна класифікувати на дві групи:**

* ієрархічні;
* неієрархічні.

Кожна з груп включає численні підходи і алгоритми.

Застосовуючи різні методи кластерного аналізу, аналітик може отримати різні рішення для однакових даних про хлібобулочну продукцію. Це вважається звичайним явищем.

**Основними завданнями кластерного аналізу є:**

* Створення типології або класифікації досліджуваних видів хлібобулочних виробів;
* Дослідження та визначення прийнятих концептуальних схем групування продукції;
* Висунення гіпотез на підставі результатів дослідження даних про асортимент;
* Перевірка гіпотез про те, чи справді типи продукції, які були виділені певним чином, мають місце в наявних даних.

**Кластерний аналіз потребує здійснення таких послідовних кроків:**

1. Проведення вибірки видів хлібобулочної продукції для кластеризації;
2. Визначення множини ознак, за якими будуть оцінюватися відібрані продукти;
3. Оцінка міри подібності видів продукції;
4. Застосування кластерного аналізу для створення груп подібних хлібобулочних виробів;
5. Перевірка достовірності результатів кластерного рішення.

## Представлення результатів застосування обраного методу

Для проведення кластерного аналізу використовувався програмний продукт IBM SPSS Modeler, адаптований для аналізу даних хлібопекарського виробництва.

IBM SPSS Modeler - програмне забезпечення для інтелектуального аналізу даних, що поєднує в собі всі необхідні технічні та аналітичні інструменти для щоденної роботи з даними про виробництво хлібобулочних виробів, розробки та впровадження ефективних прогностичних моделей для управління асортиментом.

Економічний ефект від використання IBM SPSS Modeler в рішенні задач аналізу хлібопекарського виробництва забезпечується кількома складовими:

простота доступу до даних про продукцію, їх обробки, агрегування і зміни структури;

швидка побудова та оцінка якості моделей на основі використання ефективних статистичних методів і алгоритмів машинного навчання;

оперативне впровадження побудованих моделей в практику прийняття рішень щодо асортименту;

швидке отримання віддачі від інвестицій в програмне і апаратне забезпечення за рахунок високої продуктивності, інтегрованості і масштабованості;

швидкість і ефективність процесу аналізу даних про хлібобулочну продукцію.

Процес моделювання полягає у виявленні в даних про асортимент стійких закономірностей, які можуть бути використані для прийняття управлінських рішень.

У IBM SPSS Modeler реалізована трирівнева архітектура обробки даних про хлібопекарське виробництво.

Завдання, які не потребують великих обсягів обчислень і доступу до великих масивів даних про продукцію, можуть виконуватися в IBM SPSS Modeler на локальних робочих станціях.

Тепер перейдемо до кластерного аналізу бізнес-процесу «Автоматизація виробництва хлібобулочних виробів на підприємстві ТОВ «Одеський коровай»».

Аналіз будемо проводити за допомогою метода «К-середніх». Даний метод належить до неієрархічних методів, також він є одним із найшвидших для кластерного аналізу виробничих даних.

Алгоритм k-середніх будує k кластерів хлібобулочної продукції, розташованих на максимально великих відстанях один від одного. Основний тип задач, які вирішує алгоритм k-середніх - наявність припущень (гіпотез) щодо числа кластерів продукції, при цьому вони повинні бути різні настільки, наскільки це можливо. Вибір числа k може базуватися на результатах попередніх досліджень асортименту, теоретичних міркуваннях або інтуїції.

Загальна ідея алгоритму: задане фіксоване число k кластерів, види хлібобулочної продукції зіставляються кластерам так, що середні в кластері (для всіх змінних) максимально можливо відрізняються один від одного.

Вхідними даними для кластерного аналізу хлібобулочної продукції будуть: собівартість(грн.), продажна ціна(грн.), кількість продажів(шт.), вартість продажів(грн.), загальна собівартість(грн.), прибуток(грн.) для асортименту хлібобулочних виробів підприємства. Характеристики (вартість продажів, кількість продажів, прибуток) - це показники, які отримало підприємство ТОВ «Одеський коровай» протягом 30 днів своєї виробничої діяльності, відносно всіх клієнтів.

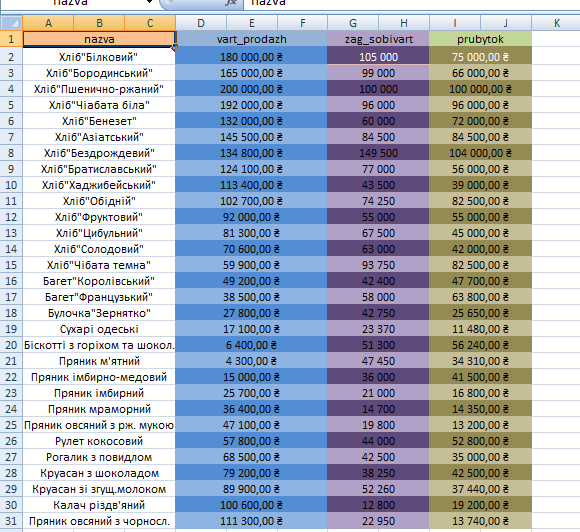


Рис.8.2.1- Представлення даних в файлі Excel

Створена модель потоку.

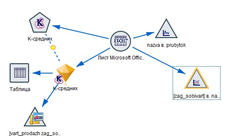
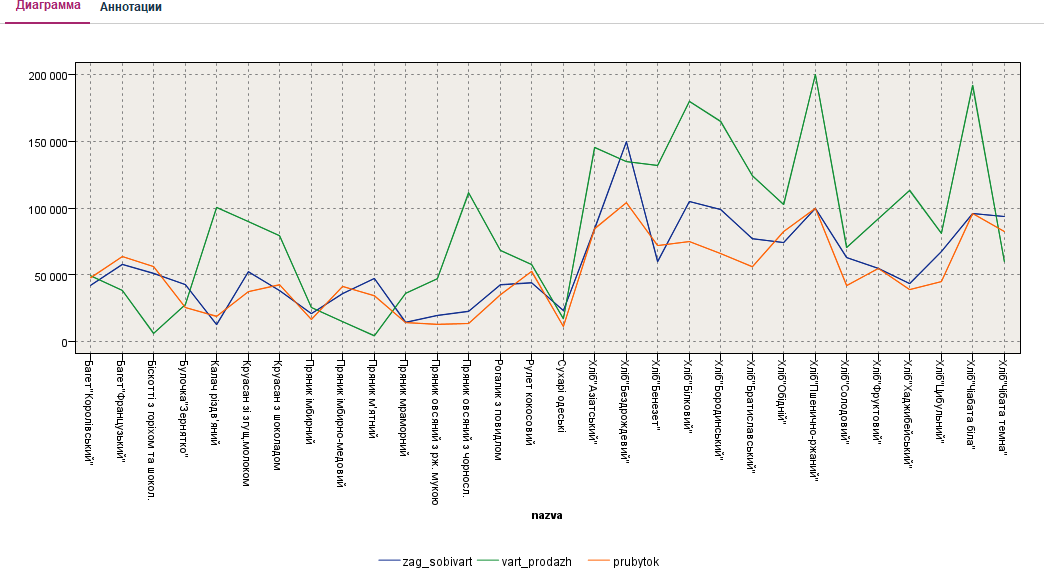


Рис.8.2.2 – Створена модель потоку

Створена модель потоку для аналізу асортименту хлібобулочних виробів.

Виконаємо побудову мультіграфіку. Покажемо характеристику( собівартість) для кожного найменування продукції. По осі Х ми вибиремо поле “nazva” файла Excel,а для по осі Y – поле «zag\_sobivartist», «prubytok»,

«zag\_prodazh».Таким чином, ми побачимо, як змінюються вибрані характеристики для кожної продукції.

Рис.8.2.3 – Мультіграфік на основі отриманих даних

Виконаємо побудову діаграми розсіювання за полем «Прибуток».Після отримання графіка зможемо проаналізувати, яка продукція приносить підприємству найбільший прибуток.

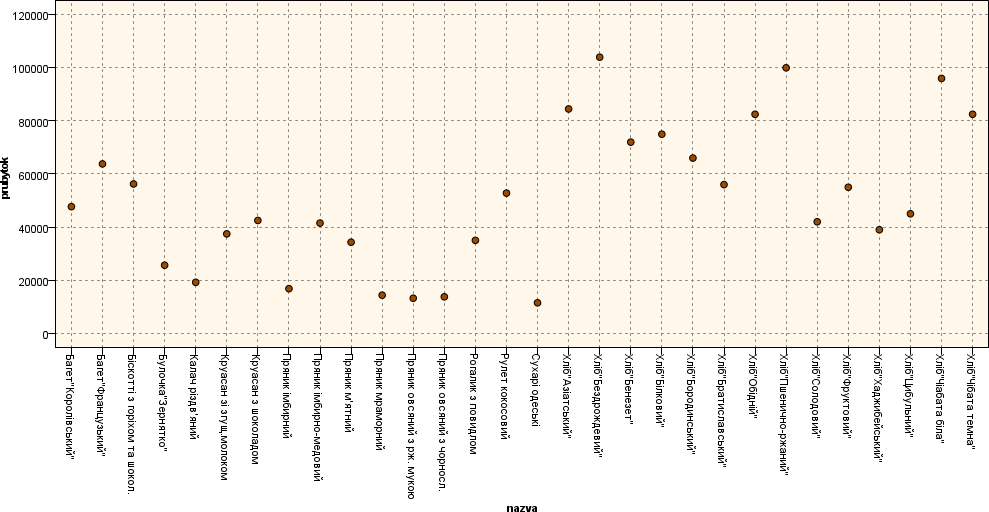


Рис.8.2.4 – Діаграма розсіювання

Проведення налаштування методу К-середніх для кластерного аналізу. В ході проведення експерименту з визначення кількості кластерів, було визначено, що їх оптимальна кількість становить 6, так як при більшій їх кількості різниця між кластерами стає дуже малою.

Результати.



**6**

**3**

Побудуємо діаграму «Круги частот» для того, щоб показати розміри отриманих кластерів асортименту хлібобулочних виробів.

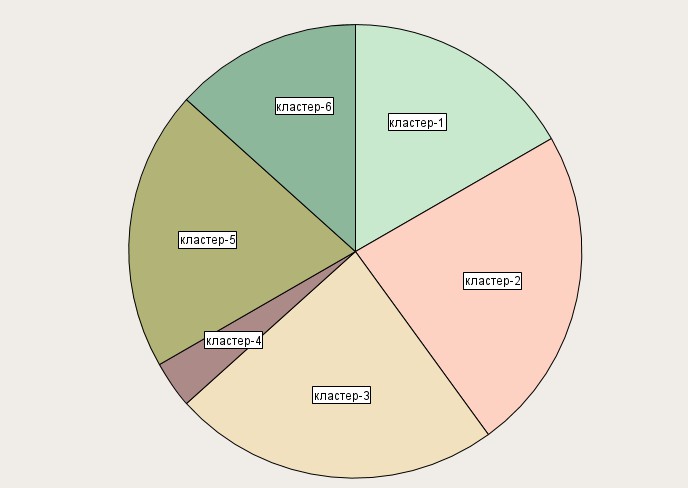


Рис.8.2.5 – Розміри кластерів В відсотковому відношенні отримаємо такі дані:

* кластер «1» (Преміум хліб) - 14%
* кластер «2» (Базовий асортимент) - 20%
* кластер «3» (Масовий сегмент) - 34%
* кластер «4» (Спеціальні вироби) - 5%
* кластер «5» (Економ-сегмент) - 16%
* кластер «6» (Сезонна продукція) - 11%

Тепер покажемо результати відпрацювання методу К-середніх.

Для відображення результату кластеризації в табличному вигляді ми перейшли в робочому вікні IBM SPSS Modeler у вкладку «Висновок» і перетягнули в робочу область блок «Таблиця», отриману модель кластеризації підключили до цього блоку. Відкривши блок «Таблиця», натиснули «Застосувати», «Запуск» і результати проведення кластеризації відобразимо в таблиці.

Для відображення результатів кластеризації асортименту хлібобулочних виробів в графічному вигляді, перейдемо в вкладку «Діаграми», виберемо блок «Панель вибору діаграм» і підключимо модель кластеризації до цього блоку. Відкривши блок, і затиснувши **ctrl**, обираємо поля за якими буде проводитись кластеризація продукції і вказуємо тип діаграми (для нашого випадку кластеризація хлібобулочних виробів проводиться по 3 полях, тому обираємо тривимірна діаграма розсіювання). Переходимо у вкладку «Деталі» і в полі «Колір» задаємо результат кластеризації.

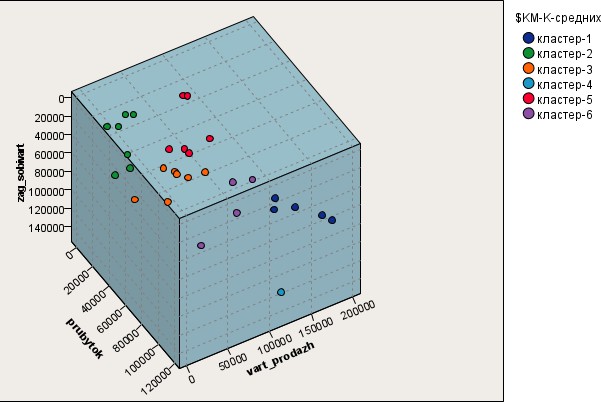


Рис.8.2.6 – Тривимірна діаграма розсіювання

**Після перегляду результатів кластеризації ми можемо охарактеризувати кожен кластер хлібобулочної продукції:**

| **№кластеру** | **Собівартість** | **Прибуток** | **Вартість продажів** |
| --- | --- | --- | --- |
| **1** | Високу | Високий | Найбільшу |
| **2** | Малу | Найменший | Найменшу |
| **3** | Середню | Середній | Невелику |
| **4** | Найбільшу | Найбільший | Середню |
| **5** | Найменшу | Середній | Малу |
| **6** | Трохи більше середньої | Трохи більше середнього | Трохи менше середньої |

Як бачимо, після аналізу характеристик кластерів хлібобулочної продукції, високий і найбільший прибуток підприємству «приносять» вироби кластерів «1» і «4», але з діаграми ми бачимо, що це кластери з найменшою кількістю об'єктів, це означає, що дуже мала кількість видів хлібобулочної продукції дає найбільший прибуток. Тому нам потрібно прагнути збільшити кількість об'єктів, які би належали до кластерів «1» і «4». Це можливо зробити, якщо, наприклад, збільшити вартість продажів певних позицій хлібобулочних виробів, що дасть можливість при тій самій кількості продажів отримувати більший прибуток.

## Розробка концепції алгоритму для автоматизації підтримки прийняття рішень

На сьогоднішній день в умовах складного динамічного середовища, що характеризується постійною невизначеністю та мінливістю політичних, економічних і соціальних факторів, основою успішного функціонування хлібопекарських підприємств є ухвалення адекватних управлінських рішень щодо виробництва та збуту хлібобулочної продукції. Сучасні системи підтримки прийняття рішення для хлібопекарського виробництва є системами, максимально пристосованими до вирішення задач повсякденної управлінської діяльності, і є інструментом, покликаним надати допомогу особам, що приймають рішення щодо асортименту та обсягів виробництва.

**Останнім часом найбільш значущими є такі проблеми у процесі розробки управлінських рішень для хлібопекарського виробництва:**

зросла взаємозалежність рішень, що приймаються в різних сферах виробництва хлібобулочних виробів;

з'явилися нові, надзвичайно складні технологічні процеси виробництва, які вимагають прийняття відповідальних рішень;

збільшилася кількість факторів, які необхідно враховувати при оцінюванні наслідків альтернатив управління асортиментом;

збільшилася кількість критеріїв оцінки альтернатив розвитку хлібопекарського виробництва.

Прийняття управлінських рішень у хлібопекарському виробництві можна розглядати як вибір з багатьох можливостей — альтернатив. Наприклад, вибір стратегій поведінки на ринку хлібобулочної продукції, вибір асортименту для виробництва. Таким чином, система підтримки прийняття рішень повинна допомагати менеджеру хлібопекарського підприємства робити такий вибір.

**Алгоритм прийняття рішень в СППР для хлібопекарського підприємства полягає у трьох фазах:**

**Перша фаза — дослідницька.** На цій фазі досліджуються знання та інформація з внутрішніх та зовнішніх джерел про ринок хлібобулочної продукції. В процесі дослідження розуміння проблем виробництва хлібобулочних виробів стає більш ясним і обґрунтованим, альтернативи асортименту і їх наслідки більш зрозумілими.

**Друга фаза — проектування.** На цьому етапі менеджер хлібопекарського підприємства формулює знання про результати альтернативних дій щодо виробництва хлібобулочної продукції, оцінює їх наслідки для функціонування системи. На цьому етапі менеджер може виробляти додаткові знання про ефективність різних видів хлібобулочних виробів.

**Третя фаза — вибір рішення.** На цьому етапі менеджер виконує вибір між альтернативами управління асортиментом хлібобулочної продукції з урахуванням їх наслідків, які були проаналізовані і досліджені на першій і другій фазах.

Прийняте рішення щодо виробництва хлібобулочних виробів може бути застосовано не одразу, але через деякий час, який підтвердить правильність прийнятого рішення. Всі три фази не завжди йдуть послідовно, вони можуть перекривати одна одну і процес прийняття рішення може бути ітеративним.

**Таким чином, з метою розширення можливостей людини СППР для хлібопекарського виробництва у підсумку виконує такі функції:**

* підтримує інформаційну модель хлібопекарського виробництва та забезпечує швидкий і асоціативний доступ до її елементів;
* підтримує генерування цілей і нестандартних альтернатив управління асортиментом хлібобулочної продукції;
* зберігає знання про раніше вирішені проблеми виробництва хлібобулочних виробів та способи їхнього вирішення;
* забезпечує створення, збереження та використання формалізованих моделей управління хлібопекарським виробництвом;
* навіть за розв'язання відносно простих проблем асортименту СППР може допомогти отримати розв'язок швидше і/або ефективніший;
* стимулює вивчення і аналіз проблем хлібопекарського виробництва особою, що приймає рішення, завдяки використанню СППР;
* забезпечує конкурентну перевагу для хлібопекарського підприємства завдяки покращеній внутрішній продуктивності управління виробництвом хлібобулочної продукції.