ІНТЕРВАЛЬНІ ОЦІНКИ ПАРАМЕТРІВ РОЗПОДІЛУ

Звіт

Роботу виконав

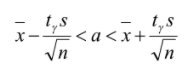
Студент групи ІТ-92

Щур Антон

**Мета роботи:** ознайомитись з методикою визначення інтервальних оцінок параметрів розподілу; дослідити, що впливає на якість інтервальних оцінок.

**Завдання №1**: Визначити інтервальні оцінки для матиматичного сподівання та середньоквадратичного відхилення при рівні довіри Рдов = 0.95

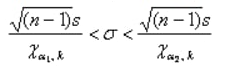
За вибіркою об’ємом 300 елементів з першої лабораторної роботи я визначив інтервальні оцінки для матиматичного сподівання. Для цього я використав формулу :



Отже, для моєї вибірки я отримав такі значення для матиматичного сподівання:



Далі я розрахував інтервальну оцінку для СКВ за формулою:

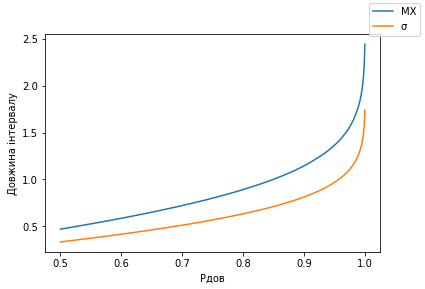


Отже, для моєї вибірки я отримав такі значення для СКВ:



**Завдання №2**: Дослідити залежність оцінок від рівня довіри.

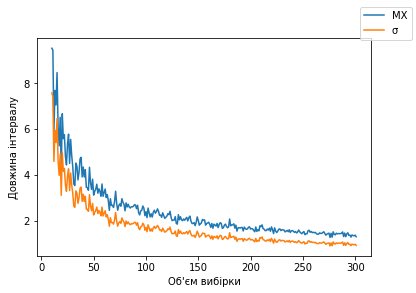
Для дослідження я зафіксував об’єм вибірки на 300 елементах, тобто на кожному кроці я використовував всю вибірку. Я провів дослідження починаючи з 0.5 і до 1(не включно) з кроком у 0.005. Після досліду я отримав такі результати, вони зображені на графіку нижче.



З даного графіку видно, що чим рівень довіри більший, тим більша довжина довірчого інтервалу. Це відповідає здоровому глузду, адже, ми не можемо гарантувати з великою точністью досить малий інтервал. Тому, якщо ми хочемо гарантувати входження оцінки у заданий інтервал з досить високою точністю(0.999), нас треба буде пожертвувати довжиною інтервалу.

**Завдання №2**: Дослідити залежність оцінок від об’єму вибірки.

Для виконання цього дослідження я зафіксував рівень довіри на значенні Рдов = 0.95 і змінював вже об’єм вибірки, при чому варіанти потрапляли у вибірку намання. Я генерував випадкову вибірку, починаючи з об’єма в 10 елементів і до 300 з кроком в 1. Після виконання досліду я отримав результати, що представив у вигляді графіка.

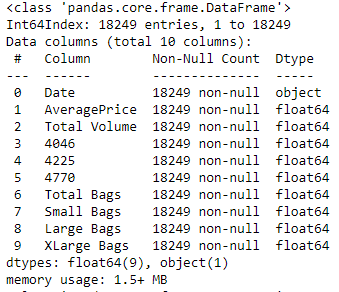


З даного графіка видно, що чим більший об’єм вибірки, тим менший інтервал оцінки. Це логічно, тому що більший об’єм вибірки дає більше інформації про саму вибірку.

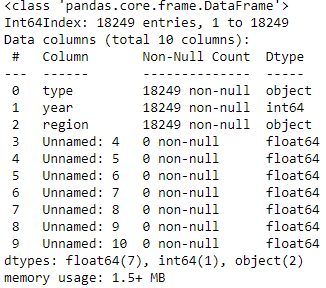
Додаткове завдання

№1 Завантажимо датасети та об’єднаємо їх в один датасет:

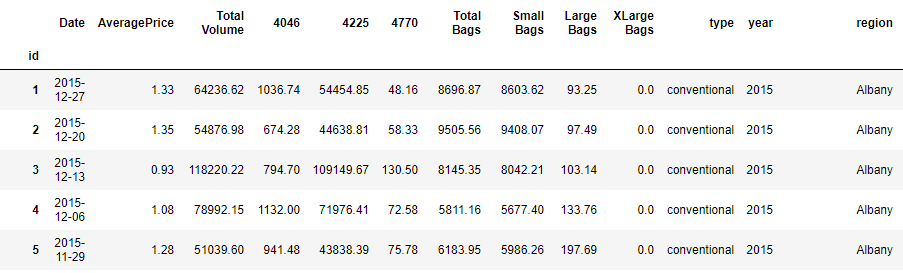
Перша половина датасету:



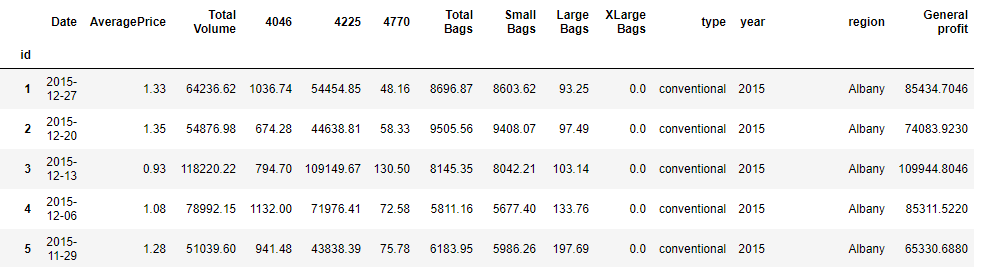
Друга половина датасету, бачимо велику кількість пустих стовпців:



Загальний датасет, як бачимо все правильно об’єдналося, пропусків немає:



№2 Створимо сповпчик з загальним прибутком:



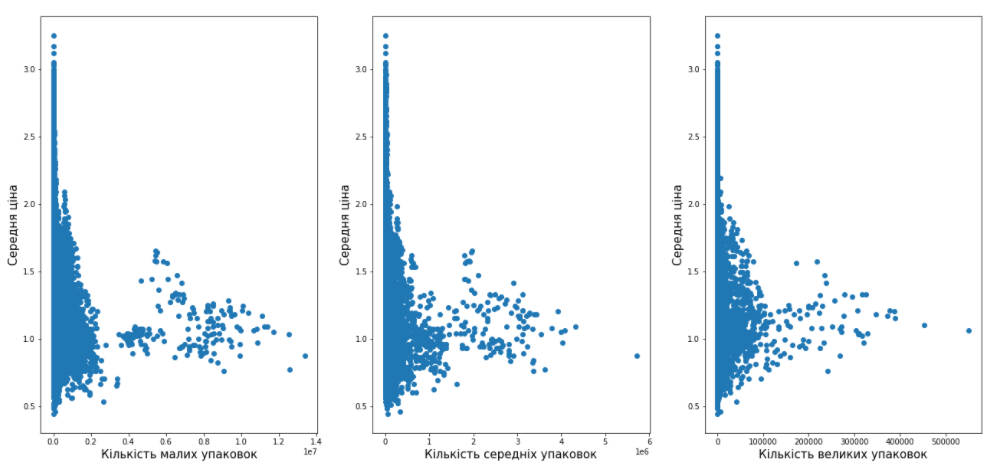
№3 Знайти загальний прибуток по органічному та неорганічному авокадо.



№4 Який рік був найбільш успішним?

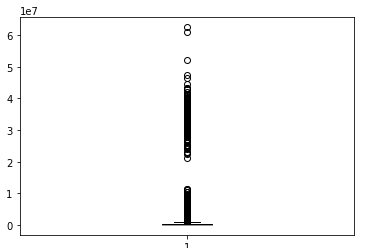
Найбільш успішним був 2017 рік

№5 Побудувати 3 графіки залежностей середньої ціни від кількості упаковок різних розмірів. Чи є очевидна залежність?

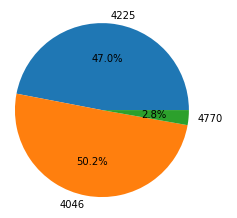


Очевидної залежності я тут не побачив.

№6 Чи є викиди в об’сягах продаж – так є



№7 Ознайомитись з функцією pie(). Побудувати кругову діаграму по кількості проданих авокадо видів 4046, 4225, 4770 у 2016 році.

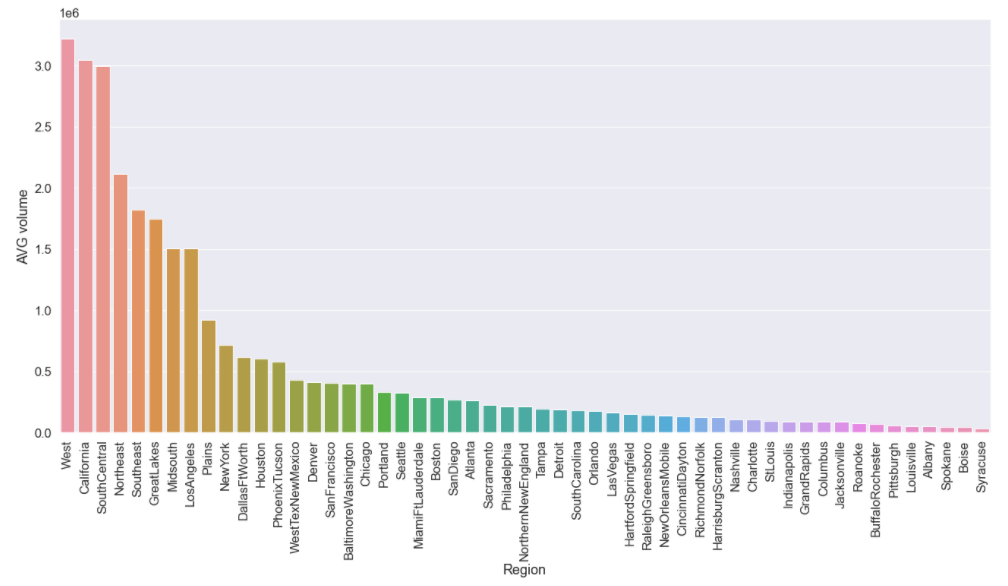


№8 В якому штаті середня ціна за весь час була мінімальною, а в якому максимальною?

Максимальною в : HartfordSpringfield

Мінімальною в : Houston

№9 Які регіони схожі по продажам авокадо? Поясніть свою відповідь.



На мою думку, такі регіони схожі:

Midsouth та Los Angeles

SanDiego та Albania

Оскільки об’єм продажу авокадо у них однаковий