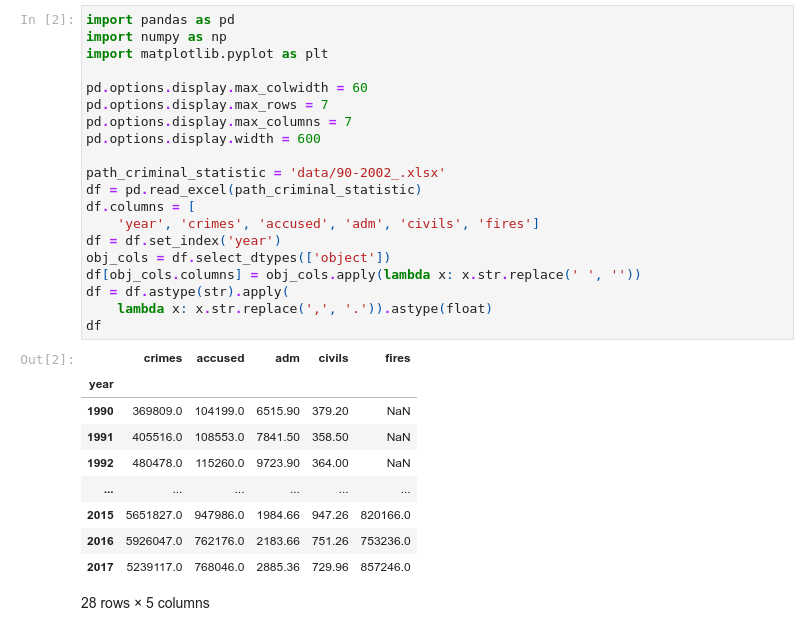
Виконання

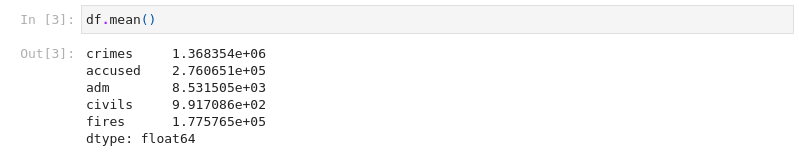
Обрати та завантажити дані

Імпортуємо бібліотеки, визначимо налаштування для pandas, завантажимо датафрейм

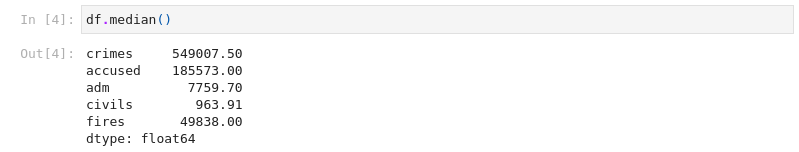
  
Рисунок 3.1 - Імпортуємо бібліотеки, визначимо налаштування для pandas, завантажимо датафрейм

Знайти математичне сподівання, медіану, моду, дисперсію, середньоквадратичне відхилення

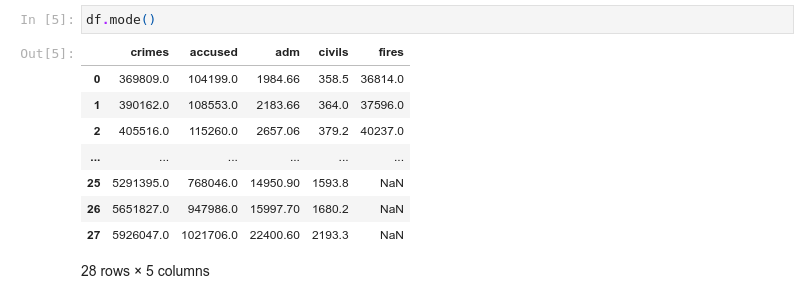
Математичне сподівання – сума всіх значень, поділена на їхню кількість.

  
Рисунок 3.2 - Математичне сподівання – сума всіх значень, поділена на їхню кількість.

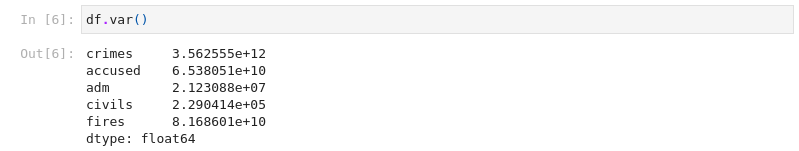
Медіана – це середина відсортованого набору даних.

  
Рисунок 3.3 - Медіана – це середина відсортованого набору даних.

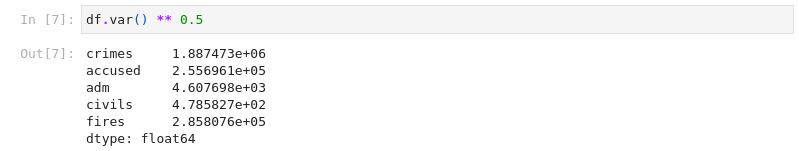
Мода – це найчастіше значення, що зустрічається.

  
Рисунок 3.4 - Мода – це найчастіше значення, що зустрічається.

Дисперсія – це середнє арифметичне квадратів різниці кожного значення з математичним сподіванням.

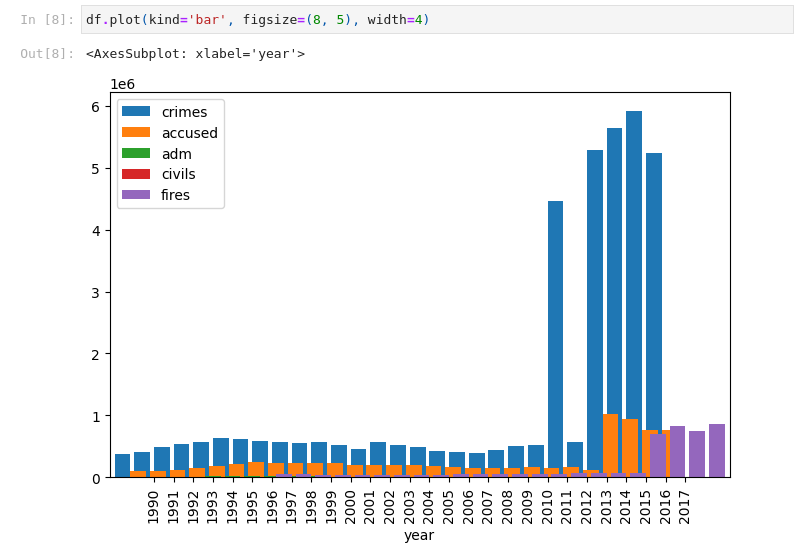
  
Рисунок 3.5 - Дисперсія – це середнє арифметичне квадратів різниці кожного значення з математичним сподіванням.

Середньоквадратичне відхилення – це дисперсія під коренем

  
Рисунок 3.6 - Середньоквадратичне відхилення – це дисперсія під коренем

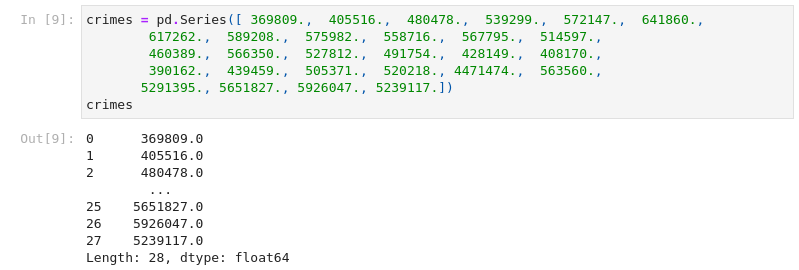
Візуалізувати завантажені дані за допомогою гістограми

Візуалізуємо дані за допомогою методу plot та оберемо тип "bar" для відображення гістограми.

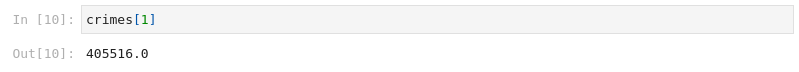
  
Рисунок 3.7 - Візуалізуємо дані за допомогою методу plot та оберемо тип "bar" для відображення гістограми.

Проробити всі дії з пункту про Series та DataFrame

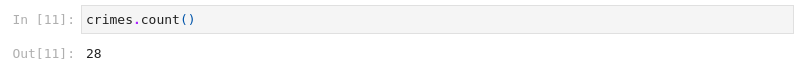
Створюємо Series зі значеннями кількості злочинів на рік.

  
Рисунок 3.8 - Створюємо Series зі значеннями кількості злочинів на рік.

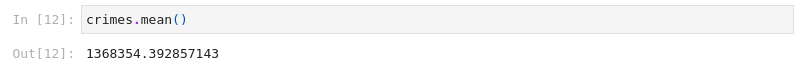
Отримуємо другий запис.

  
Рисунок 3.9 - Отримуємо другий запис.

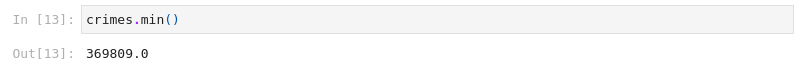
Рахуємо кількість записів.

  
Рисунок 3.10 - Рахуємо кількість записів.

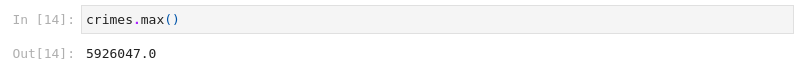
Рахуємо математичне сподівання

  
Рисунок 3.11 - Рахуємо математичне сподівання

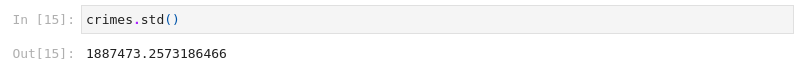
Рахуємо найменше значення

  
Рисунок 3.12 - Рахуємо найменше значення

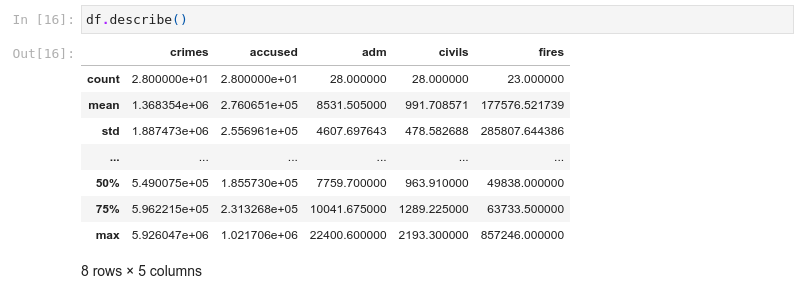
Рахуємо найбільше значення

  
Рисунок 3.13 - Рахуємо найбільше значення

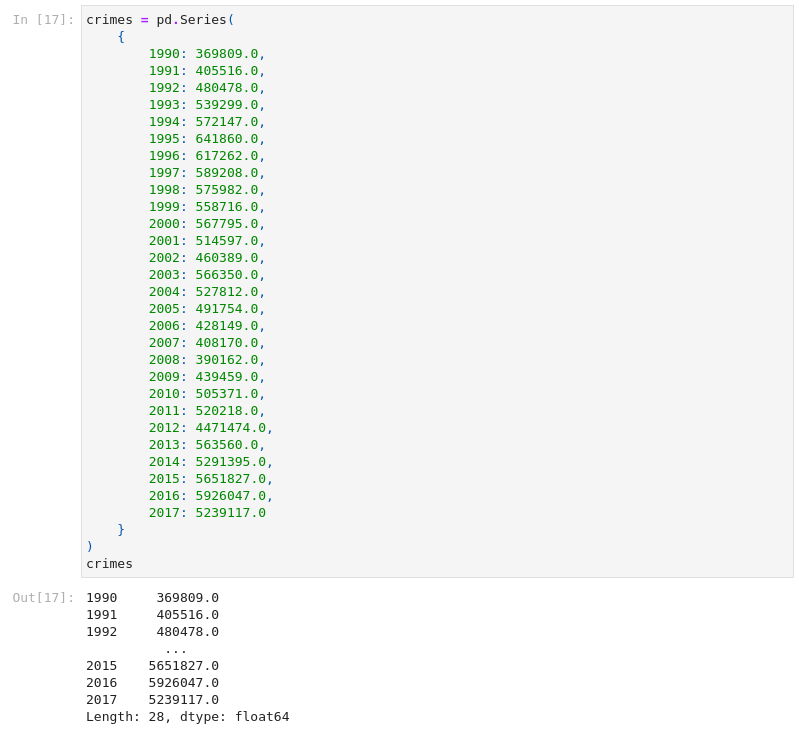
Рахуємо середньоквадратичне відхилення

  
Рисунок 3.14 - Рахуємо середньоквадратичне відхилення

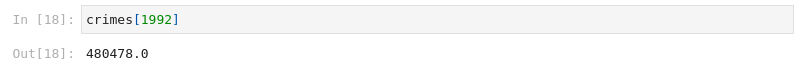
Отримуємо статистику набору даних (кількість, математичне сподівання, середньоквадратичне відхилення, найменше та найбільше значення, квантилі)

  
Рисунок 3.15 - Отримуємо статистику набору даних (кількість, математичне сподівання, середньоквадратичне відхилення, найменше та найбільше значення, квантилі)

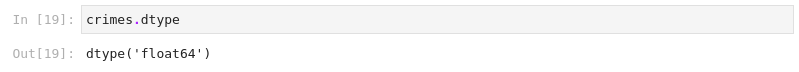
Створення Series зі словника, де ключі - роки, значення - кількість злочинів

  
Рисунок 3.16 - Створення Series зі словника, де ключі - роки, значення - кількість злочинів

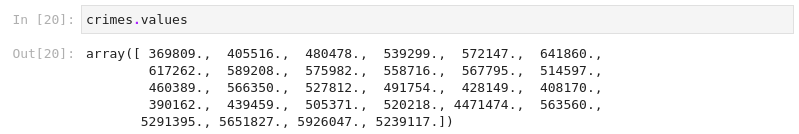
Отримуємо кількість злочинів за 1992 рік

  
Рисунок 3.17 - Отримуємо кількість злочинів за 1992 рік

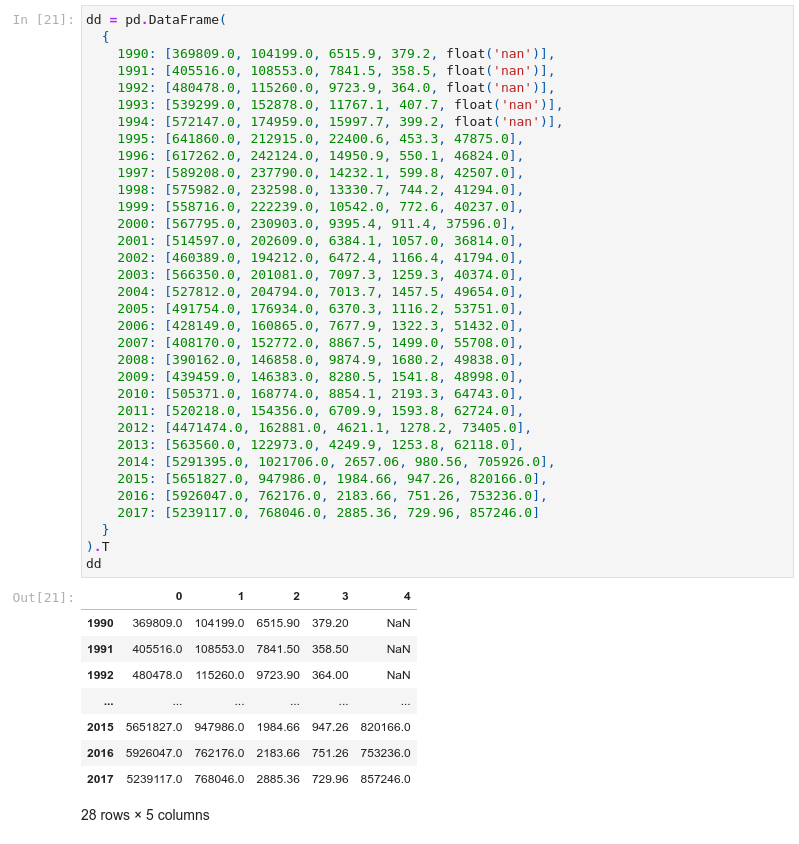
Отримуємо тип даних

  
Рисунок 3.18 - Отримуємо тип даних

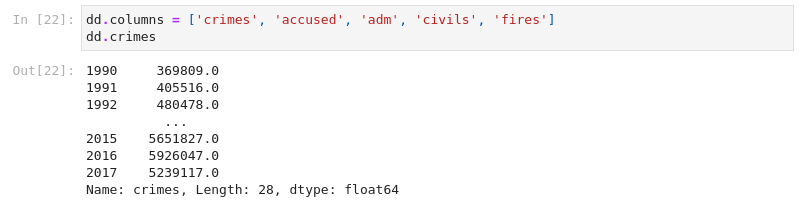
Отримуємо значення набору даних

  
Рисунок 3.19 - Отримуємо значення набору даних

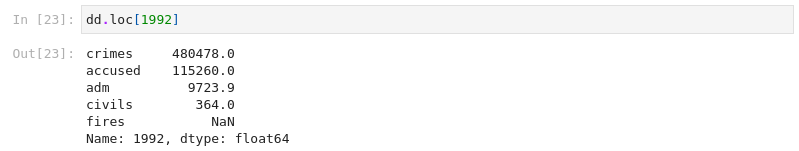
Створення DataFrame зі словника

  
Рисунок 3.20 - Створення DataFrame зі словника

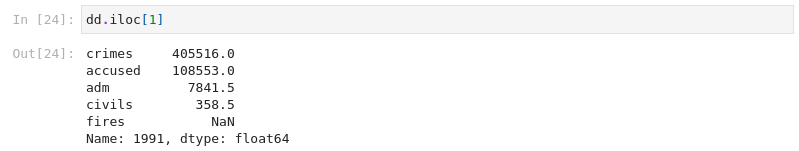
Отримаємо дані про злочини за роки незалежності, використовуючи стовпець

  
Рисунок 3.21 - Отримаємо дані про злочини за роки незалежності, використовуючи стовпець

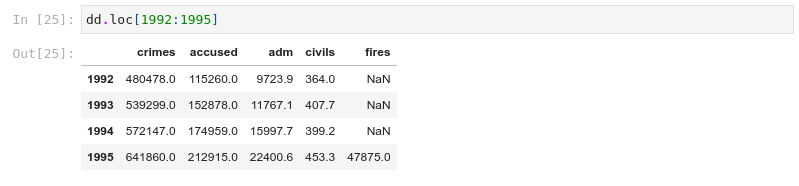
Отримаємо дані про злочини за 1992 рік, використовуючи явний індекс

  
Рисунок 3.22 - Отримаємо дані про злочини за 1992 рік, використовуючи явний індекс

Отримаємо дані про злочини за 1992 рік, використовуючи неявний індекс

  
Рисунок 3.23 - Отримаємо дані про злочини за 1992 рік, використовуючи неявний індекс

Отримуємо дані про злочини з 1992 до 1995 для явних індексів, де для явних останній вказаний індекс включається

  
Рисунок 3.24 - Отримуємо дані про злочини з 1992 до 1995 для явних індексів, де для явних останній вказаний індекс включається

Отримуємо дані про злочини з 1992 до 1994 для НЕявних індексів, де для НЕявних останній вказаний індекс НЕ включається

  
Рисунок 3.25 - Отримуємо дані про злочини з 1992 до 1994 для НЕявних індексів, де для НЕявних останній вказаний індекс НЕ включається

Отримуємо дані про злочини з 1992 І з 1995 для явних індексів

  
Рисунок 3.26 - Отримуємо дані про злочини з 1992 І з 1995 для явних індексів

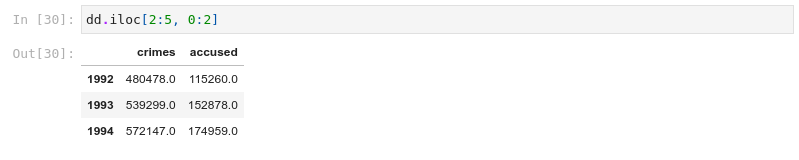
Отримуємо дані про злочини з 1992 І з 1995 для неявних індексів

  
Рисунок 3.27 - Отримуємо дані про злочини з 1992 І з 1995 для неявних індексів

Отримуємо дані про злочини явних ідексів та стовпців

  
Рисунок 3.28 - Отримуємо дані про злочини явних ідексів та стовпців

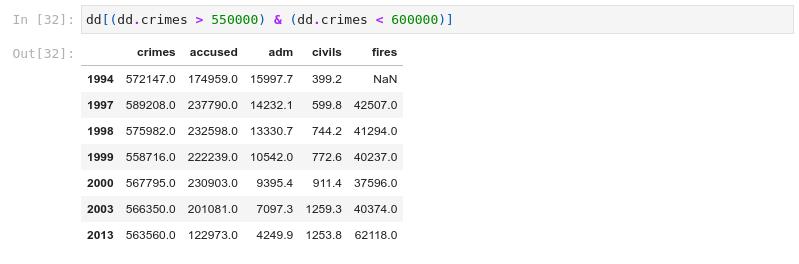
Отримуємо дані про злочини неявних ідексів та стовпців

  
Рисунок 3.29 - Отримуємо дані про злочини неявних ідексів та стовпців

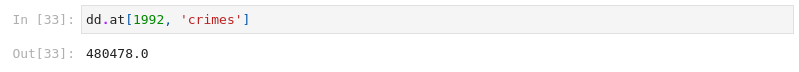
Відображаємо рядки, де кількість злочинів > 550 000

  
Рисунок 3.30 - Відображаємо рядки, де кількість злочинів > 550 000

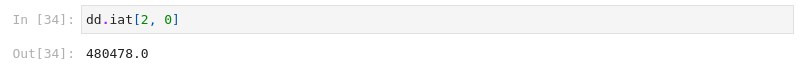
Відображаємо рядки, де кількість злочинів > 550 000 і < 600 000

  
Рисунок 3.31 - Відображаємо рядки, де кількість злочинів > 550 000 і < 600 000

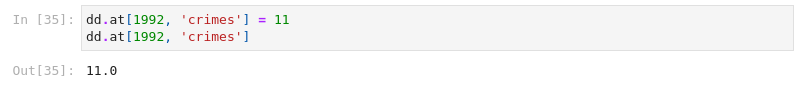
Отримуємо кількість злочинів за 1992 рік явно

  
Рисунок 3.32 - Отримуємо кількість злочинів за 1992 рік явно

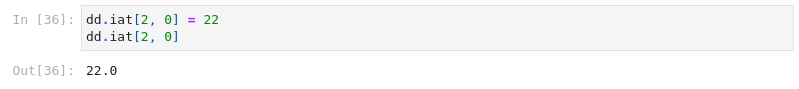
Отримуємо кількість злочинів за 1992 рік неявно

  
Рисунок 3.33 - Отримуємо кількість злочинів за 1992 рік неявно

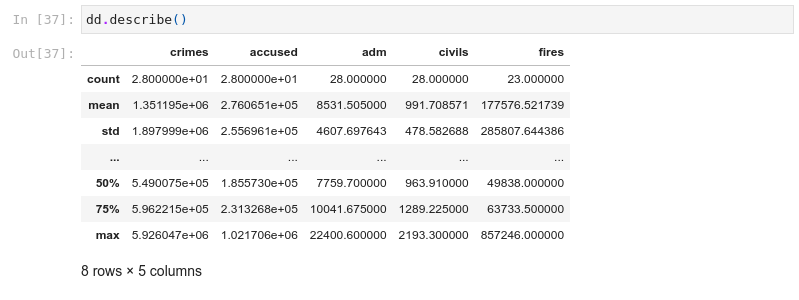
Змінюємо значення кількості злочинів за 1992 явно

  
Рисунок 3.34 - Змінюємо значення кількості злочинів за 1992 явно

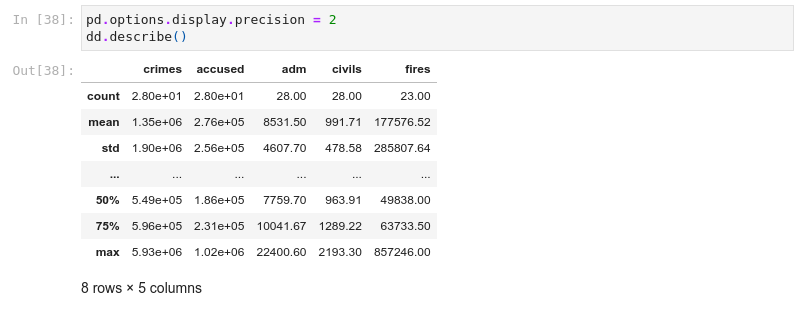
Змінюємо значення кількостізлочинів за 1992 НЕявно

  
Рисунок 3.35 - Змінюємо значення кількостізлочинів за 1992 НЕявно

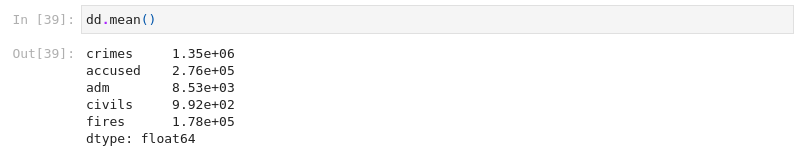
Інформацію про набір даних( кількість рядків, математичне сподівання, середньоквадратичне відхилення, найменше та найбільше значення, квантилі )

  
Рисунок 3.36 - Інформацію про набір даних( кількість рядків, математичне сподівання, середньоквадратичне відхилення, найменше та найбільше значення, квантилі )

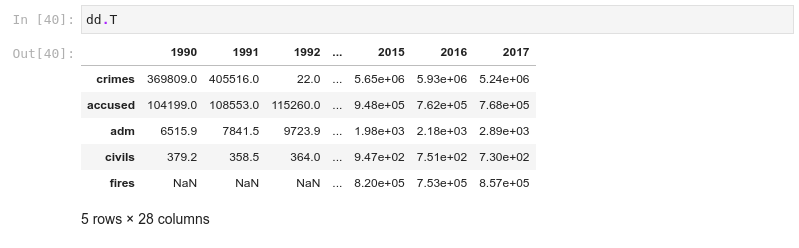
Зміна точності виведення дійсних чисел

  
Рисунок 3.37 - Зміна точності виведення дійсних чисел

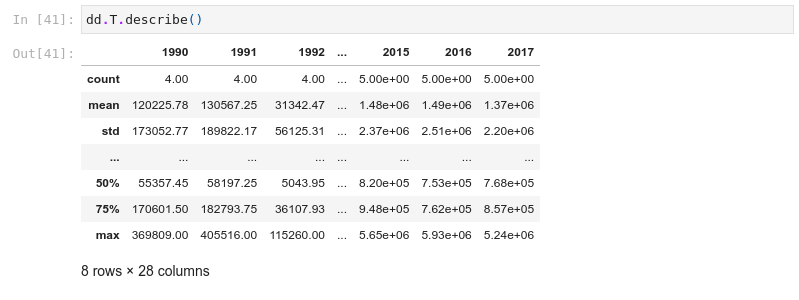
Знайдемо математичне сподівання для кожного стовпчика

  
Рисунок 3.38 - Знайдемо математичне сподівання для кожного стовпчика

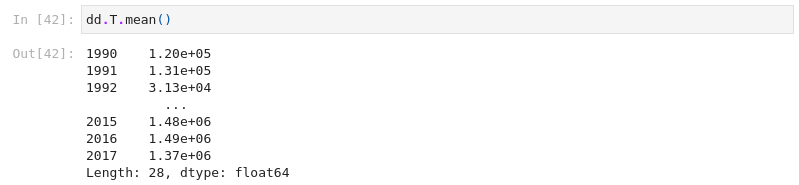
Транспонуємо набір даних

  
Рисунок 3.39 - Транспонуємо набір даних

Знаходження інформації про транспонований набір даних

  
Рисунок 3.40 - Знаходження інформації про транспонований набір даних

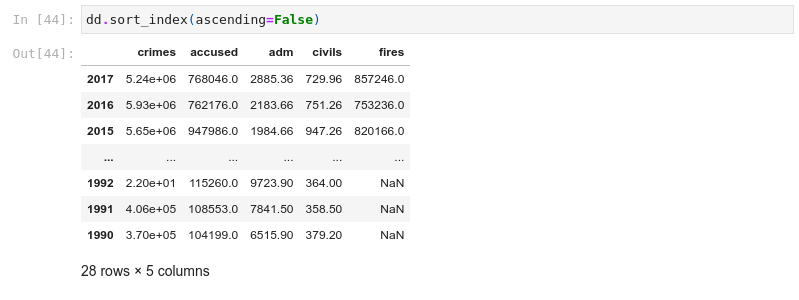
Отримуємо математичне сподівання для транспонованого набору даних, де тепер значення отримаємо не по стобцях, а по роках

  
Рисунок 3.41 - Отримуємо математичне сподівання для транспонованого набору даних, де тепер значення отримаємо не по стобцях, а по роках

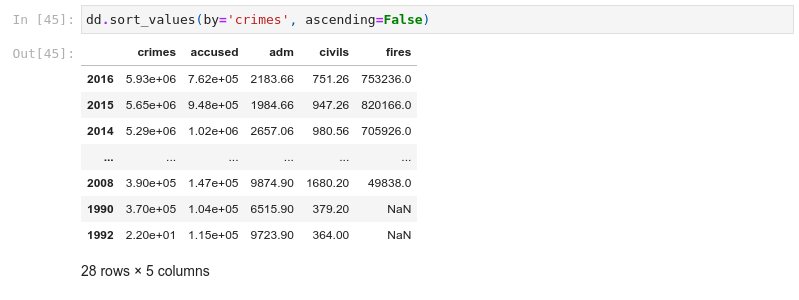
Сортуємо набір даних за індексом за зростанням

  
Рисунок 3.42 - Сортуємо набір даних за індексом за зростанням

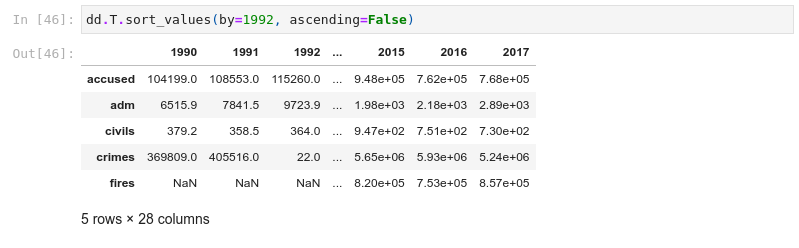
Сортуємо набір даних за індексом за спаданням

  
Рисунок 3.43 - Сортуємо набір даних за індексом за спаданням

Сортуємо стовпці за кількість скоєних злочинів

  
Рисунок 3.44 - Сортуємо стовпці за кількість скоєних злочинів

Сортуємо види порушень в транспонованому наборі даних за спаданням за 1992 рік

  
Рисунок 3.45 - Сортуємо види порушень в транспонованому наборі даних за спаданням за 1992 рік

Сортуємо стовпці за кількість скоєних злочинів, де inplace=True, щоб сортування відбувалося в самому наборі даних

  
Рисунок 3.46 - Сортуємо стовпці за кількість скоєних злочинів, де inplace=True, щоб сортування відбувалося в самому наборі даних

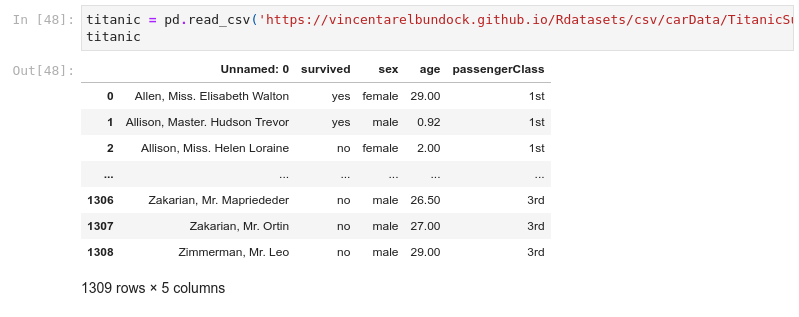
Виконати первинну обробку даних

Виконано в пункті 3.1

Прочитати набір даних катастрофи «Титаніка»

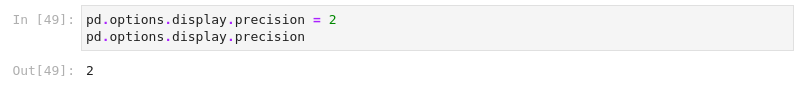
Завантажити набір даних катастрофи «Титаніка» за URL-адресою

Завантаження набору даних катастрофи «Титаніка»

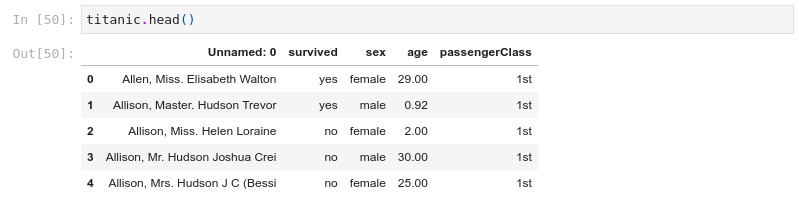
  
Рисунок 3.47 - Завантаження набору даних катастрофи «Титаніка»

Переглянути рядки набору даних катастрофи «Титаніка»

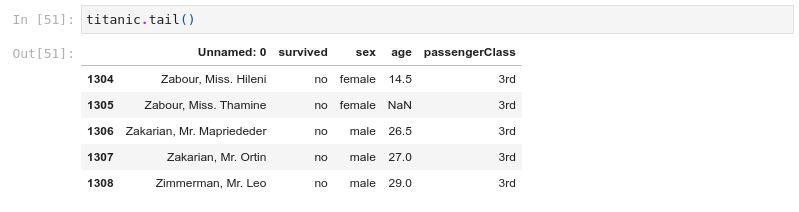
Встановлення точності виведення

  
Рисунок 3.48 - Встановлення точності виведення

Отримання перших 5-ти рядків

  
Рисунок 3.49 - Отримання перших 5-ти рядків

Отримання останніх 5-ти рядків

  
Рисунок 3.50 - Отримання останніх 5-ти рядків

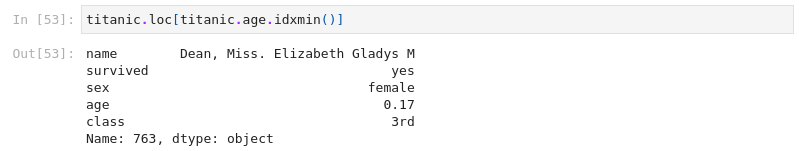
Налаштувати назви стовпців

Встановлення назви стовпців

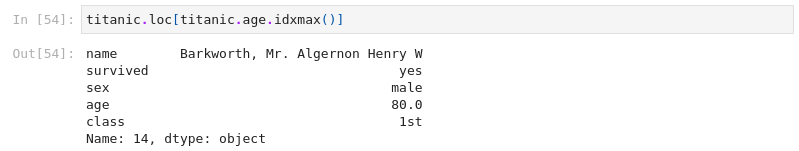
  
Рисунок 3.51 - Встановлення назви стовпців

Провести простий аналіз даних: визначити наймолодшого пасажира, найстаршого, який був середній вік пасажирів та статистику по пасажирам які вижили. Відсортувати всіх жінок з кают 1-го класу, знайти наймолодшу та найстаршу серед них, кількість вцілілих

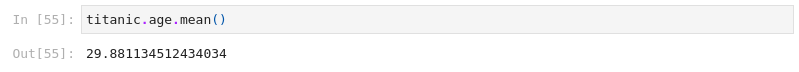
Отримання наймолодшого пасажира за допомогою знаходження рядка за його індексом, знайденого за найменшим значенням

  
Рисунок 3.52 - Отримання наймолодшого пасажира за допомогою знаходження рядка за його індексом, знайденого за найменшим значенням

Отримання найстаршого пасажира за допомогою знаходження рядка за його індексом, знайденого за найбільшим значенням

  
Рисунок 3.53 - Отримання найстаршого пасажира за допомогою знаходження рядка за його індексом, знайденого за найбільшим значенням

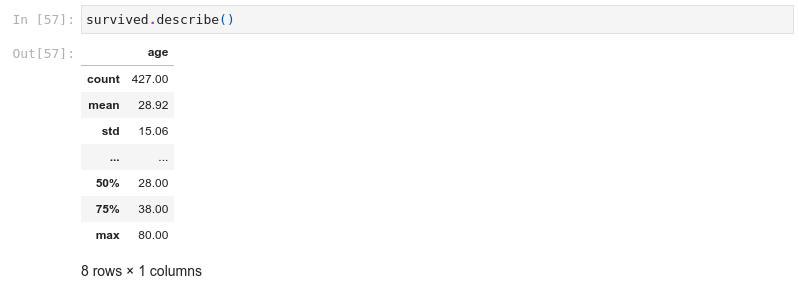
Знаходження середнього віку пасажирів методом mean

  
Рисунок 3.54 - Знаходження середнього віку пасажирів методом mean

Отримання вцілілих пасажирів за допомогою маски

  
Рисунок 3.55 - Отримання вцілілих пасажирів за допомогою маски

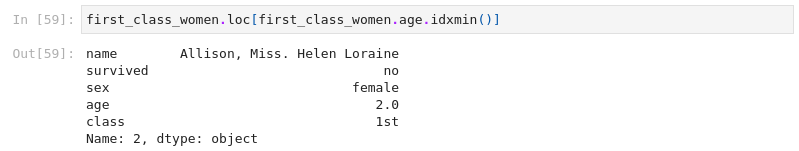
Статистика по вцілілих пасажирах (кількість, середній вік, середньоквадратичне відхилення віку, найменший вік, найбільший та квантилі)

  
Рисунок 3.56 - Статистика по вцілілих пасажирах (кількість, середній вік, середньоквадратичне відхилення віку, найменший вік, найбільший та квантилі)

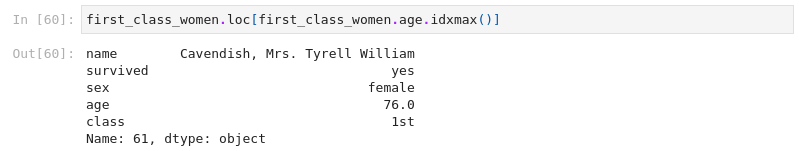
Фільтрування за статтю( жінки ) та класом ( 1ий клас )

  
Рисунок 3.57 - Фільтрування за статтю( жінки ) та класом ( 1ий клас )

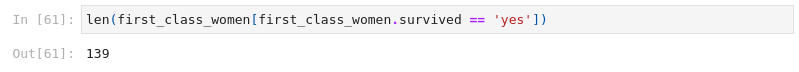
Отримання наймолодшої жінки з 1-го класу

  
Рисунок 3.58 - Отримання наймолодшої жінки з 1-го класу

Отримання найстаршої жінки з 1-ого класу

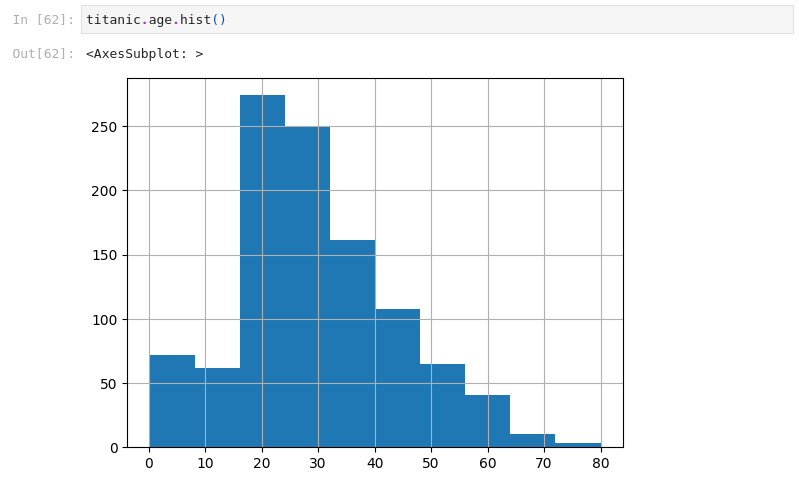
  
Рисунок 3.59 - Отримання найстаршої жінки з 1-ого класу

Знайдемо кількість жінок з першого класу, що вижили

  
Рисунок 3.60 - Знайдемо кількість жінок з першого класу, що вижили

Побудувати гістограму віку пасажирів

Гістограма віку матодом hist

  
Рисунок 3.61 - Гістограма віку матодом hist