**Для каждой переменной в вашем выводе мы получаем четыре ключевые метрики:**

* **χ²‑статистика (chi2\_stat) – измеряет «расстояние» между наблюдаемыми частотами OiO\_iOi​ и ожидаемыми EiE\_iEi​:**

**χ2=∑i=1k(Oi−Ei)2Ei. \chi^2 = \sum\_{i=1}^k\frac{(O\_i - E\_i)^2}{E\_i}.χ2=i=1∑k​Ei​(Oi​−Ei​)2​.**

* **Степени свободы (df) – равны k−1−mk - 1 - mk−1−m, где kkk – число бинов, mmm – число оценённых параметров (здесь 2: μ\muμ и σ\sigmaσ), значит df=8−1−2=5df=8-1-2=5df=8−1−2=5.**
* **p‑значение (p\_value) – вероятность получить «такие или более экстремальные» различия между OiO\_iOi​ и EiE\_iEi​, если истинно H0H\_0H0​ (данные действительно из нормального распределения с указанными параметрами).**
* **note – пометка о том, что в каком‑то бине Ei<5E\_i<5Ei​<5; у вас пусто, значит все бины удовлетворяют правилу Ei≥5E\_i\ge5Ei​≥5.**

**Критерий принятия/отвержения H0H\_0H0​**

1. **Задаём уровень значимости, обычно α=0,05\alpha=0{,}05α=0,05.**
2. **Если p\_value<αp\\_value < \alphap\_value<α, то отвергаем H0H\_0H0​ – данные не соответствуют выбранной нормальной модели.**
3. **Если p\_value≥αp\\_value \ge \alphap\_value≥α, то нет оснований отвергать H0H\_0H0​ – данные могут быть нормально распределены.**

**Ваши результаты**

| **Переменная** | **χ²‑статистика** | **df** | **p‑value** | **Интерпретация** |
| --- | --- | --- | --- | --- |
| **x1** | **1.92** | **5** | **0.8601** | **p≫0.05 нет оснований отвергнуть нормальность** |
| **x2** | **3.20** | **5** | **0.6692** | **то же** |
| **x3** | **2.944** | **5** | **0.7086** | **то же** |
| **x4** | **5.024** | **5** | **0.4130** | **то же** |
| **Y** | **1.664** | **5** | **0.8934** | **то же** |

**Все ppp-значения значительно выше 0.05. Это означает, что по χ²‑тесту у всех пяти столбцов нет статистически значимого отклонения от нормального распределения с оценёнными по выборке параметрами.**

**Что дальше?**

* **Если вам важно убедиться в нормальности более строго (особенно для малых объёмов), можно воспользоваться тестом Д’Агостино–Пирсона (scipy.stats.normaltest), Shapiro–Wilk (scipy.stats.shapiro) или другими непараметрическими критериями.**
* **Проверьте графически: Q–Q plot, гистограмма с накинутой нормальной кривой.**
* **При больших выборках даже малые отклонения могут давать значимые ppp; при малых — наоборот, тесты мало чувствительны. Всегда оценивайте результаты в контексте объёма данных и практической значимости.**

**✅ Метод K-средних:**

* Получилось **3 кластера**:
  + Кластер 0: **55 объектов**
  + Кластер 1: **53 объекта**
  + Кластер 2: **42 объекта**

📌 Это говорит о **существенной неоднородности** данных — алгоритм выявил 3 группы с различиями по признакам (X1–X4 и Y). Кластеры не равны, но близки по размеру → это нормально для реальных данных.

**✅ Метод матрицы расстояний (иерархическая кластеризация):**

* Также выделены **3 устойчивых кластера**:
  + Кластер 1: **63 объекта**
  + Кластер 2: **42 объекта**
  + Кластер 3: **45 объектов**

📌 **Дендрограмма** показывает чёткое деление на уровни — явные группы. Это ещё раз подтверждает **неоднородность совокупности**.

**✅ Вывод:**

По результатам кластерного анализа с использованием методов K-средних и иерархической кластеризации, совокупность финансовых организаций **не является однородной**. В обоих случаях данные разделяются на **три отчетливо различающихся кластера**, что подтверждает наличие значимых различий между организациями по изучаемым признакам.