

Анализ датасета «titanic.csv»

1. Рассчитать базовые статистики:

Вычислите средний возраст пассажиров, медиану, стандартное отклонение и размах. Выведите эти значения.

2. Определить выживаемость:

Рассчитайте долю выживших (столбец Survived) относительно общего числа пассажиров. Выведите результат в процентах.

3. Найти пассажиров, выживших в третьем классе, которые путешествовали без родственников

4. Распределение классов:

Постройте гистограмму распределения пассажиров по классу (Pclass): посчитайте, сколько в каждом классе было пассажиров и отобразите это графически.

5. Определить пассажиров, которые либо заплатили больше 100 за билет, либо были в первом классе, но не выжили

6. Распределение выживаемости по полу:

Постройте столбчатую диаграмму, показывающую выживаемость пассажиров (сгруппированных по полу - колонка Sex): посчитайте, сколько пассажиров выжило и визуальным образом отобразите это с учетом их пола.

Указание: изучите метод groupby() у объектов DataFrame

7. Вывести пассажиров, которые либо путешествовали с более чем двумя родственниками, либо находились в первом классе, но при этом заплатили за билет меньше 20

8. Различные зависимости:

- Построить точечный график, отображающий распределение стоимости билетов (Fare) в зависимости от возраста (Age) для выживших и погибших пассажиров. Использовать разные цвета для обозначения выживших и погибших
- Построение гистограмму средней стоимости билета (Fare) в зависимости от класса билета (Pclass)
- *** Постройте линейный график изменения стоимости билета в зависимости от возраста

9. Найти пассажиров, не входящих в возрастной диапазон от 18 до 50 лет и выживших