Т-критерий для независимых выборок позволяет определить влияет ли 2х уровневый номинальный фактор на количественный нормально распределенный по подгруппам отклик.

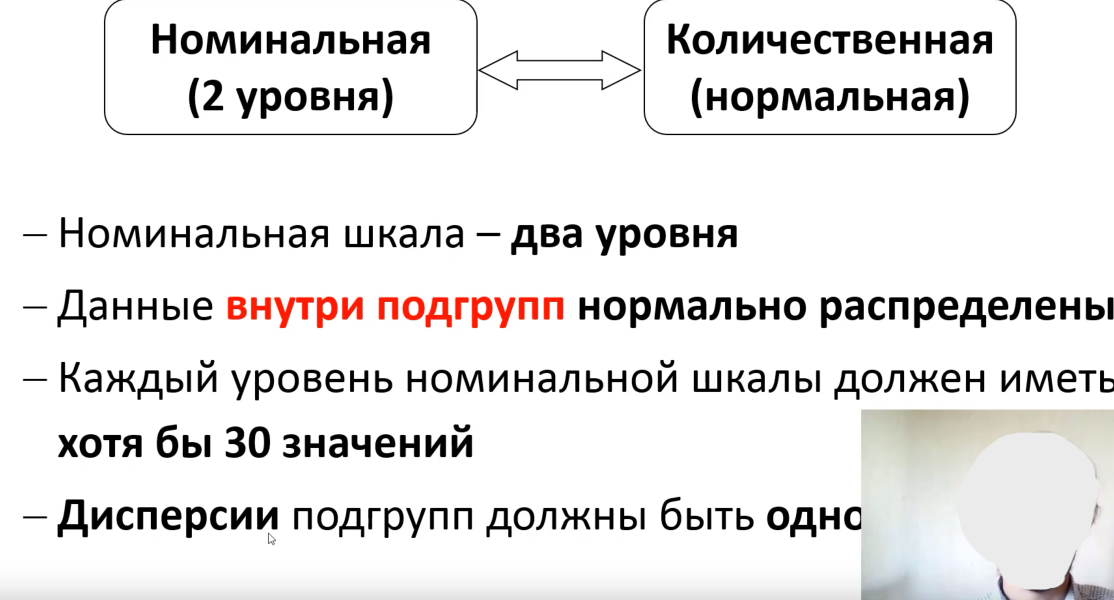
Требования к исходным данным:

+ номинальная шкала должна иметь два уровня, если больше, то пользуемся другими критериями.

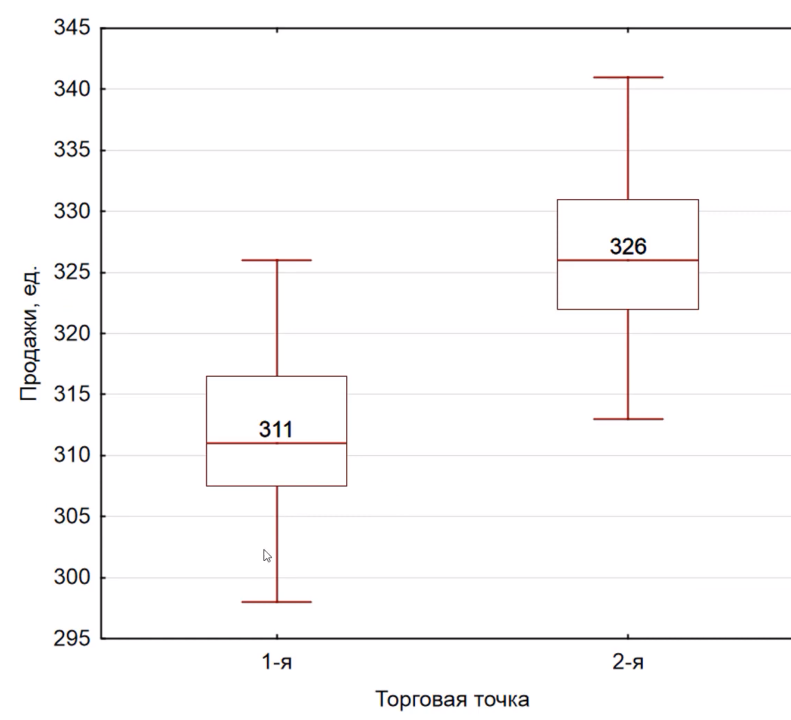
+ данные внутри подгрупп должны быть нормально распределены.

+ каждый уровень номинальной шкалы должен иметь больше 30 значений (потому что, если данных мало, то мы заранее можем предположить, что данные ненормально распределены)

+ дисперсии подгрупп должны быть однородными



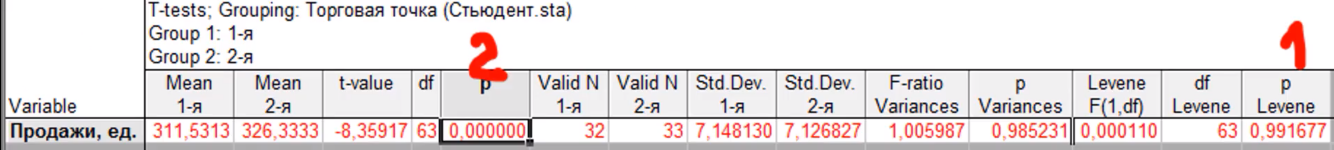
Пример:



Цель: понять, различаются ли торговые точки по продажам?

Видим, что распределение данных возможно нормально. Медианы различаются, значит гипотеза правомерна.

Применяем критерий Т-стьюдента для независимых ваборок и получаем таблицу:

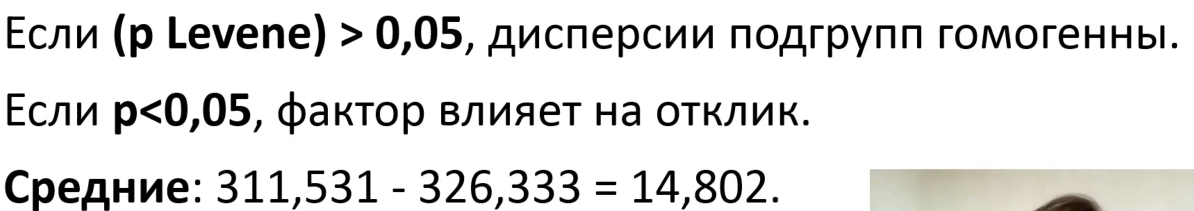


На таблице красным показаны первое и второе значение. Первое – значение теста левена, второе – т критерия стьюдента.

Тест левена: если p >0,05, то дисперсии подгрупп однородны (гомогенны). В данном случае да, однородны.

Т критерий стьюдента: если p< 0,05, то фактор влияет на отклик. В данном случае да, влияет.

Средние(разница): 311.531 - 326.333 = 14.802 Соответственно вывод: во второй точке на 14 единиц в среднем продажи выше.



**Алгоритм**

1. Проверяем подгруппы количественной шкалы на нормальность.
2. Если в обеих подгруппах распределение нормально, то проверяем однородность дисперсии с помощью теста ливеана.
3. Если дисперсии однородны, то переходим к интерпретации уровня значимости Т-критерия стьюдента, если ее нет или если ненормально распределены, то переходим к критерию U-манна уитни.
4. Если все сложилось, подгруппы распределены нормально, дисперсии однородны, то рассчитываем средние значения и интерпретируем их, то есть сравниваем средние значения и делаем вывод.

