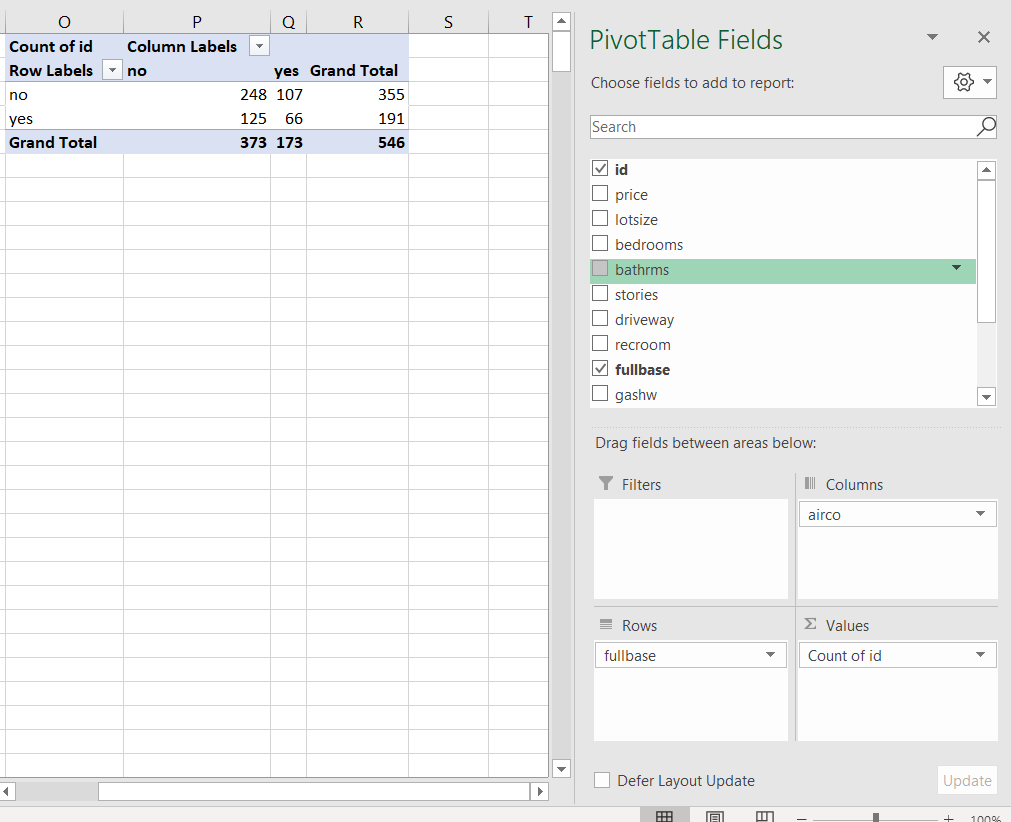
Expected values and repeated measures: demo notes

**Chi-square test of independence**

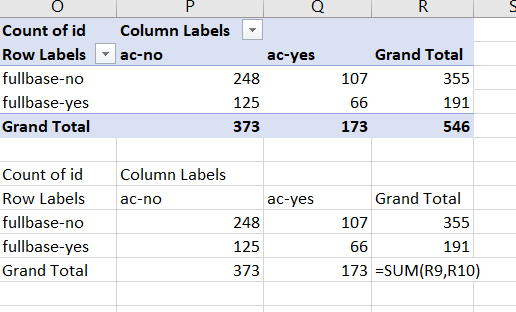
File: housing.xlsx

Is there a relationship between homes that have air conditioning and homes that have a full basement?

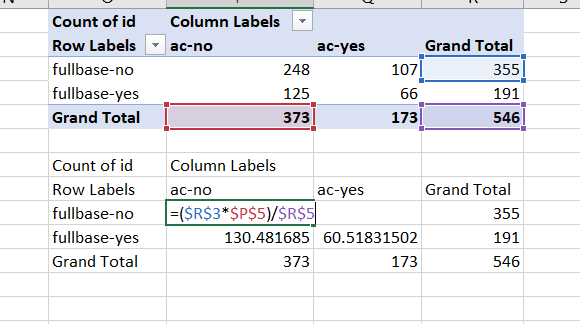
1. Insert a PivotTable from the source data. Place airco in the Columns section, fullbase in the Rows section, and count of id in the Values section.



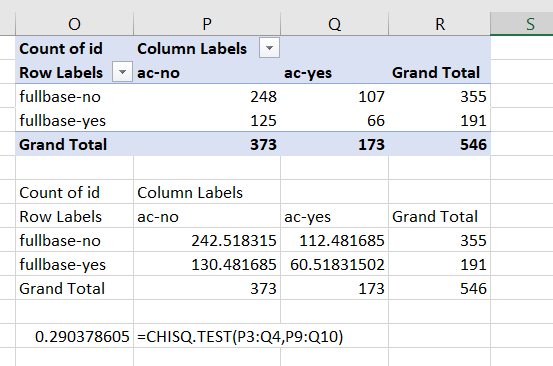
1. These are our *actual* values. We will now calculate the *expected* values based on overall proportions. Copy and paste the data from the PivotTable and add formulas to sum the row and column totals.
   1. It can also be helpful to relabel the row and column headers to know exactly which category is which.



1. We will now calculate what would be expected for each category based on a straight proportion.



1. We will now find the p-value of the chi-square test using the Excel function CHISQ.TEST(). The function takes two arguments: the first is the range of *actual* values (which come from the PivotTable), and the second from the *expected* (which we calculated).
   1. In this case, since the p-value exceeds .05, we fail to reject the null. There is no significant relationship between these categories.

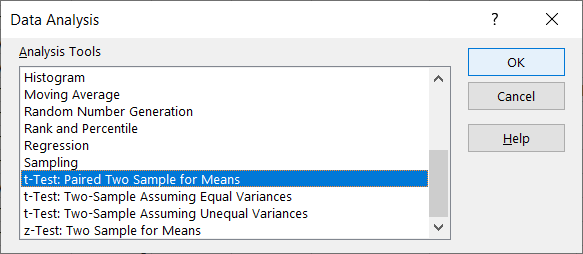


**Paired sample t-test**

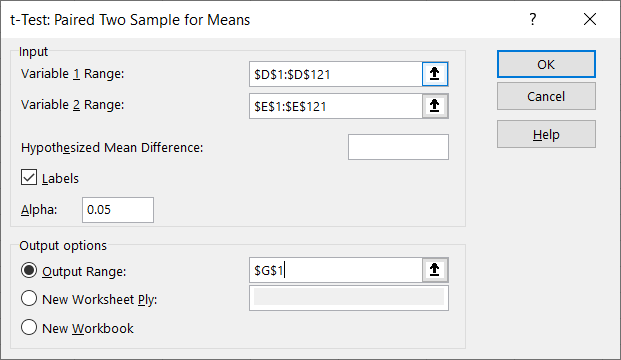
Demo: bp.xlsx

Is there a difference in blood pressures after the intervention?

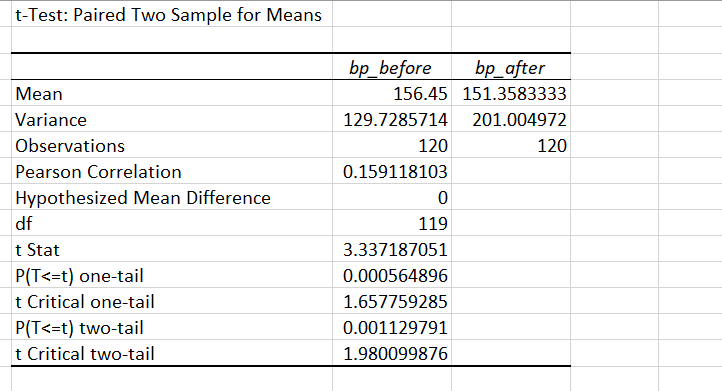
1. For this test, we will use the Data Analysis ToolPak. Find the t-test: Paired Two Sample for Means option.



1. Select the measurements for Time 1 and Time 2. The Hypothesized Mean Difference will set to zero by default, which is what we want.



1. Based on the p-value for the two-tailed t-test, we reject the null. Blood pressures are significantly lower on average after the intervention.



**Wilcoxon signed-rank test**

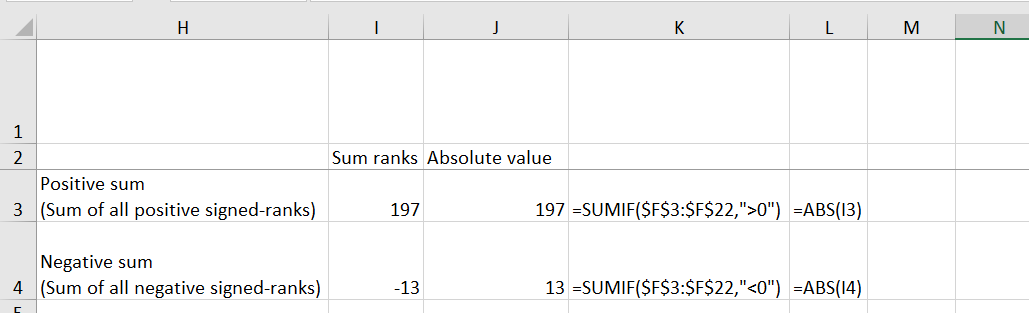
Demo: cortisol.xlsx

For this test, we will multiply the sign of each observation (+1 or -1) by the relative ranking of its absolute value, ranked ascending.

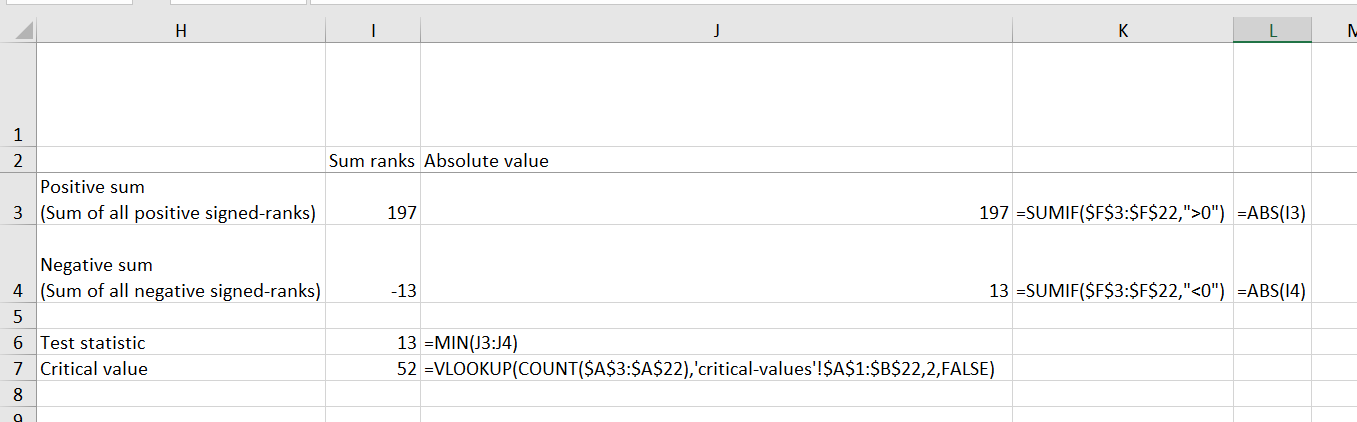
1. Enter the following formulas for columns D-F:

|  |  |  |
| --- | --- | --- |
| D | Sign | =SIGN(B3-C3) |
| E | Absolute value | =ABS(B3-C3) |
| F | Signed-rank | =RANK.AVG(E3,$E$3:$E$22,1)\*D3 |

1. Now we take the sum of all positive and negative ranks, and find their absolute values:



1. Our test statistic is whatever of these two values is smaller, and our critical value is found from a lookup of the critical-values worksheet based on our sample size.



1. If the test statistic is *less than* the critical value, we reject the null.

