

# “How can I test for a trend across a categorical variable in Stata?”

Authored by  
**stats writer**

July 1, 2024

## RECOMMENDED CITATION

stats writer (2024). “How can I test for a trend across a categorical variable in Stata?”. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=164882>

To test for a trend across a categorical variable in Stata, one can use the command "xi" followed by the variable name and the trend option. This will create a series of dummy variables representing the different categories of the variable and will allow for the analysis of a linear trend. The "xi" command can also be used with other statistical tests, such as regression or ANOVA, to examine the relationship between the categorical variable and a continuous outcome variable. Additionally, the "tabi" command can be used to create a contingency table and conduct a chi-square test to determine if there is a significant trend across the categorical variable. Overall, Stata provides various tools and commands to effectively test for trends in categorical variables and assess their significance.

## **How can I test for a trend across a categorical variable? | Stata FAQ**

**In a oneway analysis of variance (ANOVA), the alternative hypothesis is that the means across the levels of a factor variable are not equal. In the situation where the factor variable has a natural ordering, this alternative hypothesis may not be the exact hypothesis that one wishes to test. An alternative hypothesis is that the responses systematically increase or decrease over the levels of the factor variable. In Stata, the nptrend command performs a non-parametric test of trend for the ranks of across ordered groups. The test is an**

## **extension of the Wilcoxon rank-sum test.**

**For a quick example, we are going to see if there is a trend of ranks of math score over the levels of socio-economic status (ses) using the hsb2 dataset.**

**NOTE: This syntax works for Stata 17 and later.**

**use <https://stats.idre.ucla.edu/stat/stata/notes/hsb2>,  
clear  
(high school and beyond (200 cases))**

**nptrend math, group(ses) cuzick**

**Cuzick's test with rank scores**

**Number of observations = 200**

**Number of groups = 3**

**Number of response levels = 40**

---

**| Mean**

**| response Number**

**Group | Group score score of obs**

-----+-----

**ses |**

**low | 1 49.17021 47**

**middle | 2 52.21053 95**

**high | 3 56.17241 58**

-----

**Statistic = 12.18**

**Std. err. = 2.954912**

**z = 4.122**

**Prob > |z| = 0.0000**

**From the test, there appears to be a trend in math scores across the ordered levels of ses.**

**The syntax for the nptrend command changed in Stata version 17. The code above works with Stata 17; for earlier versions of Stata, the code below will work.**

**nptrend math, by(ses)**

**ses score obs sum of ranks**

**low 1 47 3637.5**

**middle 2 95 9283.5**

**high 3 58 7179**

**$z = 4.12$**

**Prob > |z| = 0.000**

ARABPSYCHOLOGY.COM