



# TRƯỜNG ĐẠI HỌC NHA TRANG

## KHOA CÔNG NGHỆ THỰC PHẨM

# HƯỚNG DẪN XỬ LÝ SỐ LIỆU THỰC NGHIỆM BẰNG SPSS

1. SỰ KHÁC BIỆT CÓ Ý NGHĨA THỐNG KÊ GIỮA CÁC GIÁ TRỊ TRUNG BÌNH
2. PHÂN TÍCH HỒI QUY

Người trình bày: Nguyễn Văn Tăng

Nha Trang - 4/2013

# VẤN ĐỀ ĐẶT RA

- Thủ tục xử lý thống kê cần áp dụng là gì?
- Phần mềm hỗ trợ như thế nào?
- Thao tác xử lý ra sao?
- Cách đọc kết quả như thế nào?

# 1. SỰ KHÁC BIỆT CÓ Ý NGHĨA THỐNG KÊ GIỮA CÁC GIÁ TRỊ TRUNG BÌNH

- Khi biểu diễn kết quả thí nghiệm

Thông tin cần có trên bảng, biểu đồ gồm:

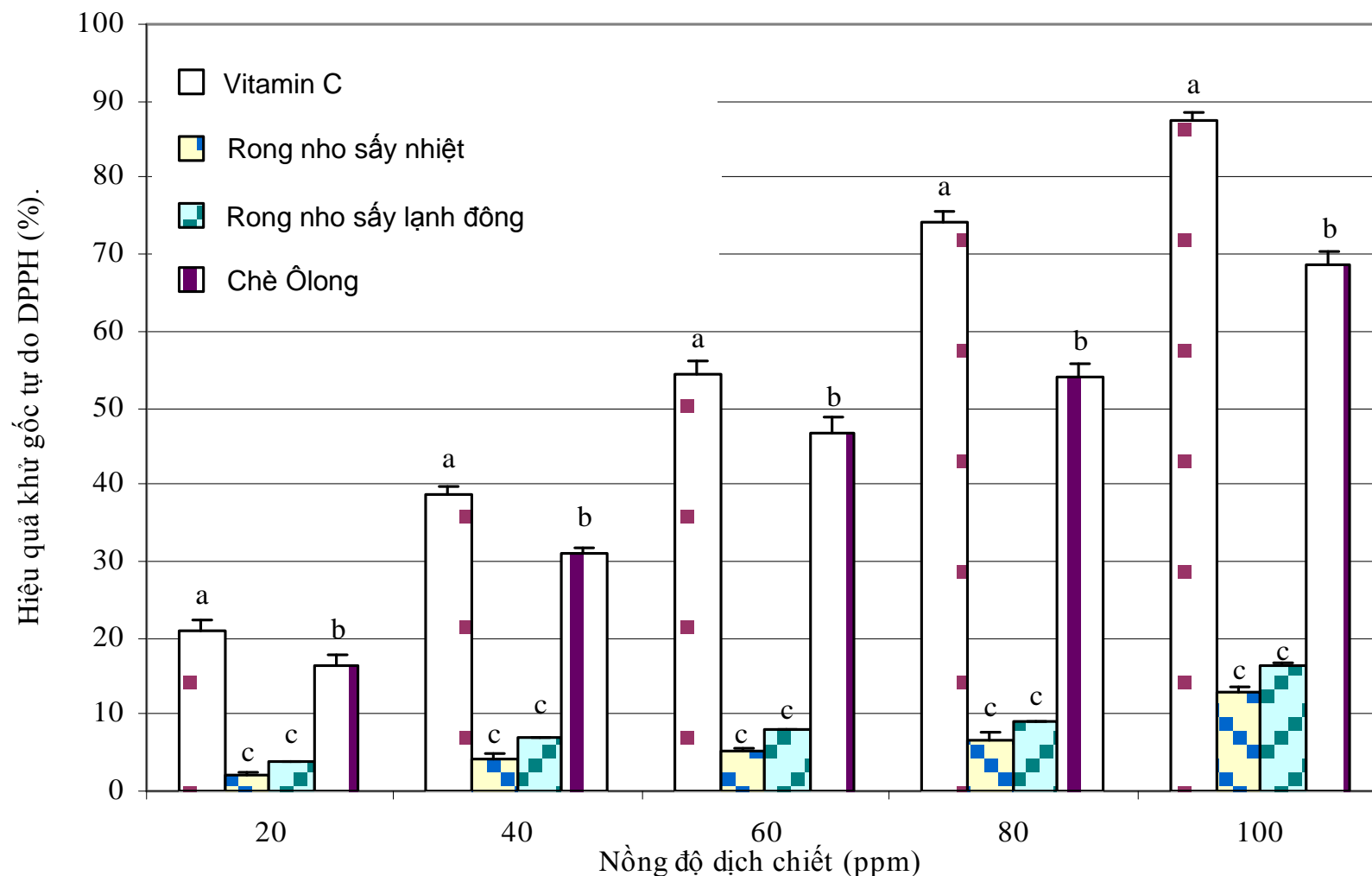
- Giá trị trung bình
- Độ lệch chuẩn
- Ký hiệu chỉ có hay không sự khác biệt có ý nghĩa thống kê ( $p < 0,05$ ) giữa các giá trị trung bình (thường dùng chữ cái a,b,c)

**Bảng 1-Năng suất sấy rong nho tươi, năng suất chiết và hàm lượng phenolic tổng số của rong nho sấy nhiệt, sấy lạnh đông và chè Ôlong**

Loại nguyên liệu	Năng suất sấy (g/100 g nguyên liệu tươi)	Năng suất chiết (g/100 g chất khô)	Hàm lượng phenolic tổng số (mg GAE*/g chất khô)
Rong nho sấy nhiệt	$2.21 \pm 0.08^{a**}$	$7.62 \pm 0.11^c$	$1.30 \pm 0.02^c$
Rong nho sấy lạnh đông	$2.26 \pm 0.09^a$	$11.02 \pm 0.05^b$	$2.04 \pm 0.03^b$
Chè Ôlong	nd	$15.99 \pm 0.09^a$	$13.58 \pm 0.01^a$

\*GAE, Gallic acid equivalent; nd, Không xác định

Nguồn: Nguyễn và cộng sự (2011)



Hình 1-Khả năng khử gốc tự do DPPH của dịch chiết rong nho sấy nhiệt và sấy lạnh đông so sánh với chè Ôlong và vitamin C (chất chuẩn).

Nguồn: Nguyễn và cộng sự (2011)

# BÀI TẬP TÌNH HUỐNG

- Khảo sát ảnh hưởng của nồng độ ethanol đến hàm lượng polyphenol trong dịch chiết chè đen
- Mục tiêu cần giải quyết:
  - Có sự khác biệt có ý nghĩa thống kê về hàm lượng polyphenol trong dịch chiết chè đen ở các nồng độ ethanol khác nhau hay không?

# Số liệu thực nghiệm

<b>Nồng độ Ethanol</b>	<b>TPC (ppm)</b>	
	<b>Lần 1</b>	<b>Lần 2</b>
50	419.730	399.636
60	507.152	480.356
70	524.320	528.264
80	602.112	617.600
90	614.528	619.520
99,5	222.870	242.604

Nguồn: Trần và Nguyễn (2013)

## ➤ Thủ tục xử lý thống kê cần áp dụng:

- Phân tích phương sai 1 yếu tố
- Yếu tố: Nồng độ ethanol với số mức  $k = 6$  (tương ứng với các nồng độ: 50; 60; 70; 80; 90; và 99,5%)
- Giá trị quan sát: Hàm lượng polyphenol (ppm)



## ➤ Phần mềm SPSS hỗ trợ xử lý:

1. Dùng One way ANOVA – chỉ ra có hay không sự khác biệt có ý nghĩa thống kê giữa các giá trị trung bình
2. Dùng One way ANOVA (Options - Homogeneity of variance test ) chỉ ra có hay không sự đồng nhất giữa các phương sai của các nhóm – là cơ sở để chọn test trong Post Hoc Multiple Comparisions

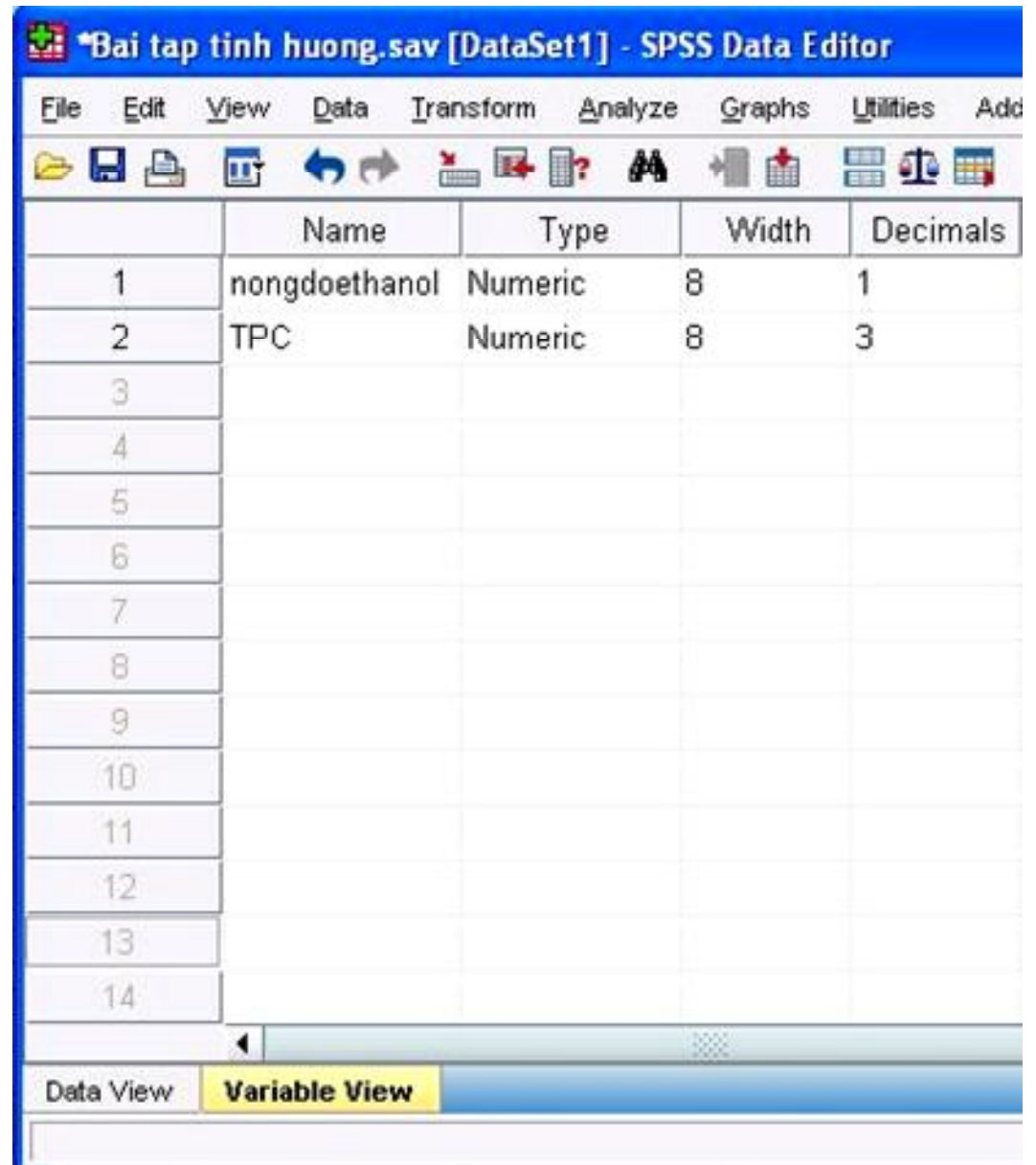
➤ **Phần mềm SPSS hỗ trợ xử lý:**

3. Dùng One way ANOVA (Post Hoc Multiple Comparisons) để chỉ ra cụ thể sự khác biệt có ý nghĩa thống kê giữa cặp giá trị trung bình nào và ở nồng độ ethanol là bao nhiêu thì hàm lượng polyphenol trong dịch chiết chè đen là cao (thấp) nhất

## ➤ Thao tác xử lý

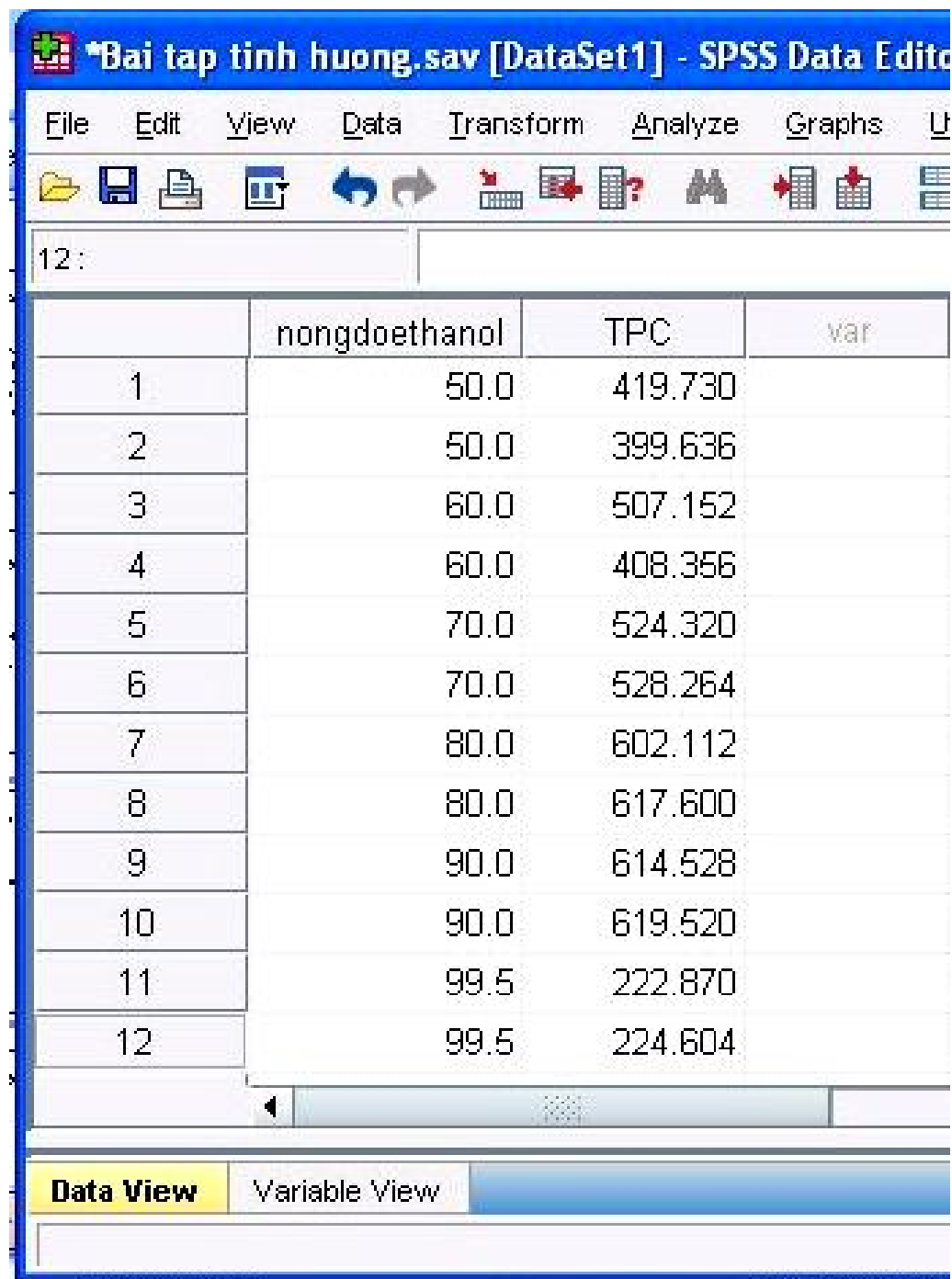
Mở giao diện  
SPSS

Khai báo tên biến:  
chọn **Variable  
View** – khai báo  
tên biến theo  
hàng



## ➤ Thao tác xử lý

Khai báo dữ  
liệu của biến:  
chọn **Data View**  
– khai báo dữ  
liệu của biến  
theo cột



\*Bai tap tinh huong.sav [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs U

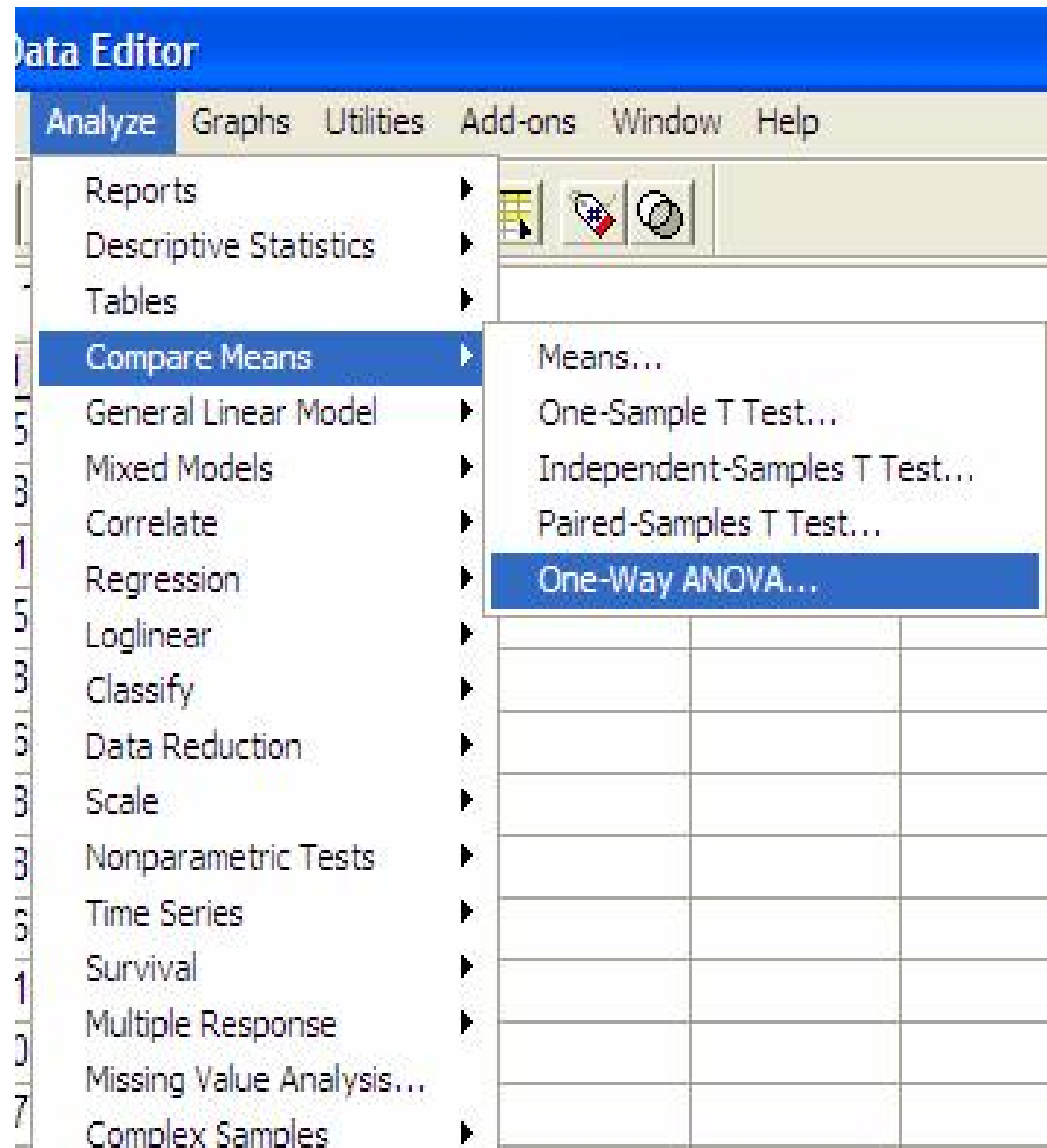
12 :

	nongdoethanol	TPC	var
1	50.0	419.730	
2	50.0	399.636	
3	60.0	507.152	
4	60.0	408.356	
5	70.0	524.320	
6	70.0	528.264	
7	80.0	602.112	
8	80.0	617.600	
9	90.0	614.528	
10	90.0	619.520	
11	99.5	222.870	
12	99.5	224.604	

Data View Variable View

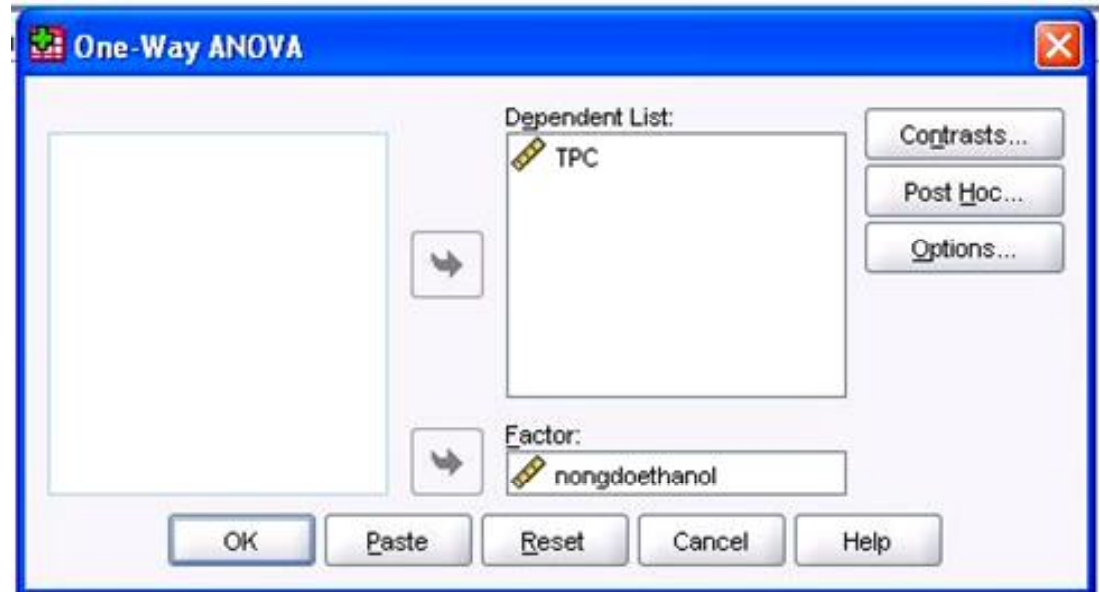
## ➤ Thao tác xử lý

Thực hiện phân  
tích dữ liệu:  
Chọn Analyze  
→ Compare  
Means → **One  
way ANOVA**



## ➤ Thao tác xử lý

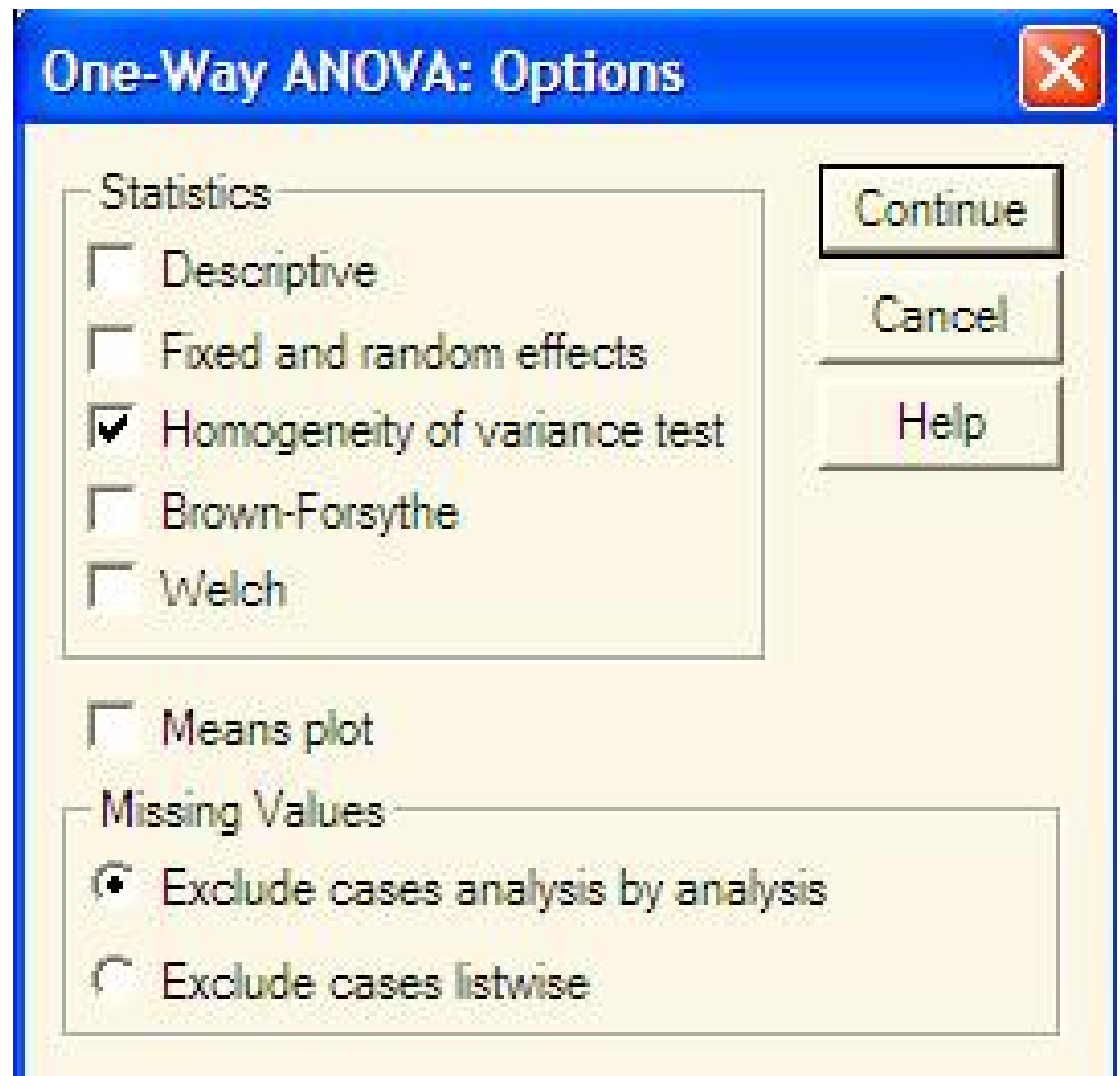
Trong hộp thoại khai báo yếu tố và biến phụ thuộc (giá trị quan sát)



## ➤ Thao tác xử lý

Chọn option →  
Đánh dấu vào:

- **Homogeneity of variance test** để biết phương sai của các nhóm có đồng nhất hay không



**One-Way ANOVA: Options**

☐ Descriptive  
☐ Fixed and random effects  
☒ Homogeneity of variance test  
☐ Brown-Forsythe  
☐ Welch

☐ Means plot

**Missing Values**  
☒ Exclude cases analysis by analysis  
☐ Exclude cases listwise

Continue  
Cancel  
Help

### Test of Homogeneity of Variances

TPC

Levene Statistic	df1	df2	Sig.
1.782E16	5	6	.000

### ANOVA

TPC

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	217358.658	5	43471.732	49.930	.000
Within Groups	5223.889	6	870.648		
Total	222582.547	11			



## ➤ Cách đọc kết quả

Nhìn vào giá trị Sig trong bảng ANOVA để biết có hay không sự khác biệt các giá trị trung bình:

$\text{Sig} \leq \alpha (0,05)$ : có sự khác biệt → tiếp tục phân tích sâu ANOVA

**ANOVA**

TPC					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	217358.658	5	43471.732	49.930	.000
Within Groups	5223.889	6	870.648		
Total	222582.547	11			

## ➤ Cách đọc kết quả

- Nhìn vào giá Sig trong bảng Test Homogeneity để biết có hay không sự khác biệt về phương sai giữa các nhóm.

**Test of Homogeneity of Variances**

TPC

Levene Statistic	df1	df2	Sig.
1.782E16	5	6	.000

## ➤ Cách đọc kết quả

- Sig >  $\alpha$ : không có sự khác biệt về phương sai → Trong Post Hoc Multiple Comparisons chọn nhóm test (LSD; Duncan; Scheffé; Tukey; Bonferroni; ...) cho phần phân tích sự khác biệt giữa các giá trị trung bình.

## ➤ Cách đọc kết quả

- $\text{Sig} \leq \alpha$ : có sự khác biệt về phương sai → Trong Post Hoc Multiple Comparisons chọn nhóm test (Tamhane's T2; Dunnett's T3; Dunnett's C...) cho phần phân tích sự khác biệt giữa các giá trị trung bình.

## ➤ Cách đọc kết quả

- Với tình huống này giá trị sig trong bảng ANOVA và Test Homogeneity đều  $< 0,05$ , nên trong Post Hoc Multiple Comparisions chọn test Dunnett's T3 cho phần phân tích sự khác biệt giữa các giá trị trung bình
- Kết quả thể hiện trong bảng sau

# Post Hoc

## Multiple Comparisons

Dependent Variable:TPC

	(I) nong doet hano	(J) nong doet hano	Mean Difference (I- J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Dunnett T3	50	60	-48.0710000	5.0409E1	.950	-1.399960E3	1303.818219
		70	-116.6090000	1.0238E1	.127	-394.662986	161.444986
		80	-2.001730E2	1.2685E1	.022	-327.479322	-72.866678
		90	-207.3410000	1.0352E1	.063	-462.157051	47.475051
		99.5	185.9460000	1.0084E1	.086	-131.453848	503.345848
	60	50	48.0710000	5.0409E1	.950	-1.303818E3	1399.960219
		70	-68.5380000	4.9437E1	.838	-1.671973E3	1534.897274
		80	-152.1020000	5.0001E1	.485	-1.599511E3	1295.307116
		90	-159.2700000	4.9461E1	.470	-1.755517E3	1436.977066
		99.5	234.0170000	4.9405E1	.333	-1.379151E3	1847.184589
	70	50	116.6090000	1.0238E1	.127	-161.444986	394.662986
		60	68.5380000	4.9437E1	.838	-1.534897E3	1671.973274
		80	-83.5640000	7.9911E0	.129	-277.805468	110.677468
		90	-90.7320000	3.1810E0	.007	-122.222273	-59.241727
		99.5	302.5550000	2.1541E0	.002	267.458197	337.651803

80	50	200.1730000'	1.2685E1	.022	72.866678	327.479322
	60	152.1020000	5.0001E1	.485	-1.295307E3	1599.511116
	70	83.5640000	7.9911E0	.129	-110.677468	277.805468
	90	-7.1680000	8.1363E0	.966	-178.416069	164.080069
	99.5	386.1190000'	7.7923E0	.031	147.031668	625.206332
90	50	207.3410000	1.0352E1	.063	-47.475051	462.157051
	60	159.2700000	4.9461E1	.470	-1.436977E3	1755.517066
	70	90.7320000'	3.1810E0	.007	59.241727	122.222273
	80	7.1680000	8.1363E0	.966	-164.080069	178.416069
	99.5	393.2870000'	2.6422E0	.004	340.660956	445.913044
99.5	50	-185.9460000	1.0084E1	.086	-503.345848	131.453848
	60	-234.0170000	4.9405E1	.333	-1.847185E3	1379.150589
	70	-3.025550E2'	2.1541E0	.002	-337.651803	-267.458197
	80	-3.861190E2'	7.7923E0	.031	-625.206332	-147.031668
	90	-3.932870E2'	2.6422E0	.004	-445.913044	-340.660956

\*. The mean difference is significant at the 0.05 level.

# Homogeneous

## TPC

	nong doet ha...	N	Subset for alpha = 0.05			
			1	2	3	4
Duncan <sup>a</sup>	99.5	2	2.2373E2			
	50	2		4.0968E2		
	60	2		4.5775E2	4.5775E2	
	70	2			5.2629E2	
	80	2				6.0985E2
	90	2				6.1702E2
	Sig.		1.000	.154	.059	.816

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 2.000.



## ➤ Cách đọc kết quả

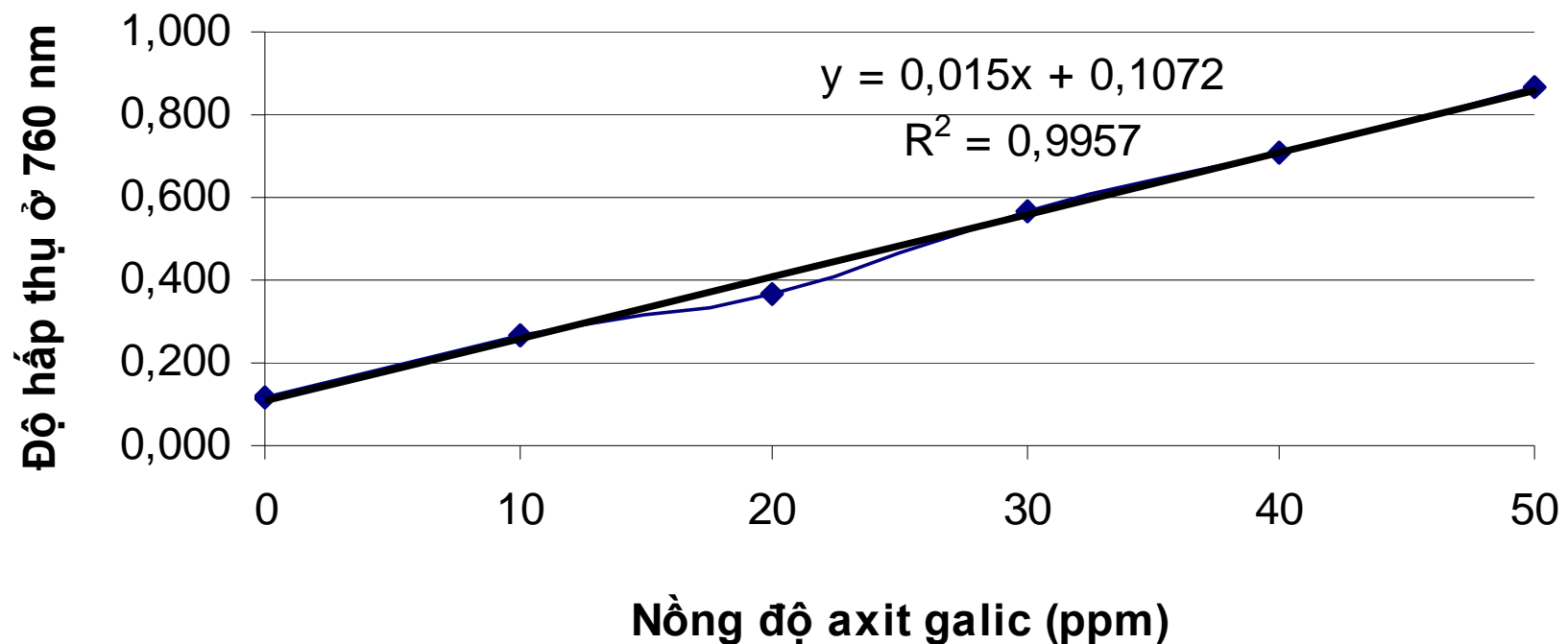
Kết quả thể hiện trong hộp Post Hoc Test:

- Dấu \* thể hiện sự khác biệt giữa các cặp giá trị trung bình
- Khi đó khi trình bày số liệu trong biểu đồ ta sẽ dùng các chữ cái a,b,c... để phân biệt sự khác nhau có ý nghĩa thống kê giữa các cặp giá trị trung bình

Nồng độ Ethanol	50	60	70	80	90	99,5
TPC	b	bc	c	d	d	a

## 2. PHÂN TÍCH HỒI QUY

- ✓ Cho phép dự đoán diễn biến của quá trình dựa vào phương trình hồi quy và hệ số xác định ( $R^2$ )



Nguồn: Trần và Nguyễn (2013)

➤ Thủ tục xử lý cần áp dụng:

- Phân tích hồi quy tuyến tính: Phương trình tuyến tính:  $y = ax + b$
- Biến độc lập: nồng độ ethanol; biến phụ thuộc: hàm lượng polyphenol

## ➤ **Phần mềm SPSS hỗ trợ xử lý:**

1. Dùng Regression - Linear

2. Dùng Regression - Curve estimation

## ➤ Thao tác xử lý

1. Chọn Analyze →  
Regression → Linear

2. Chọn Analyze →  
Regression → curve  
estimation

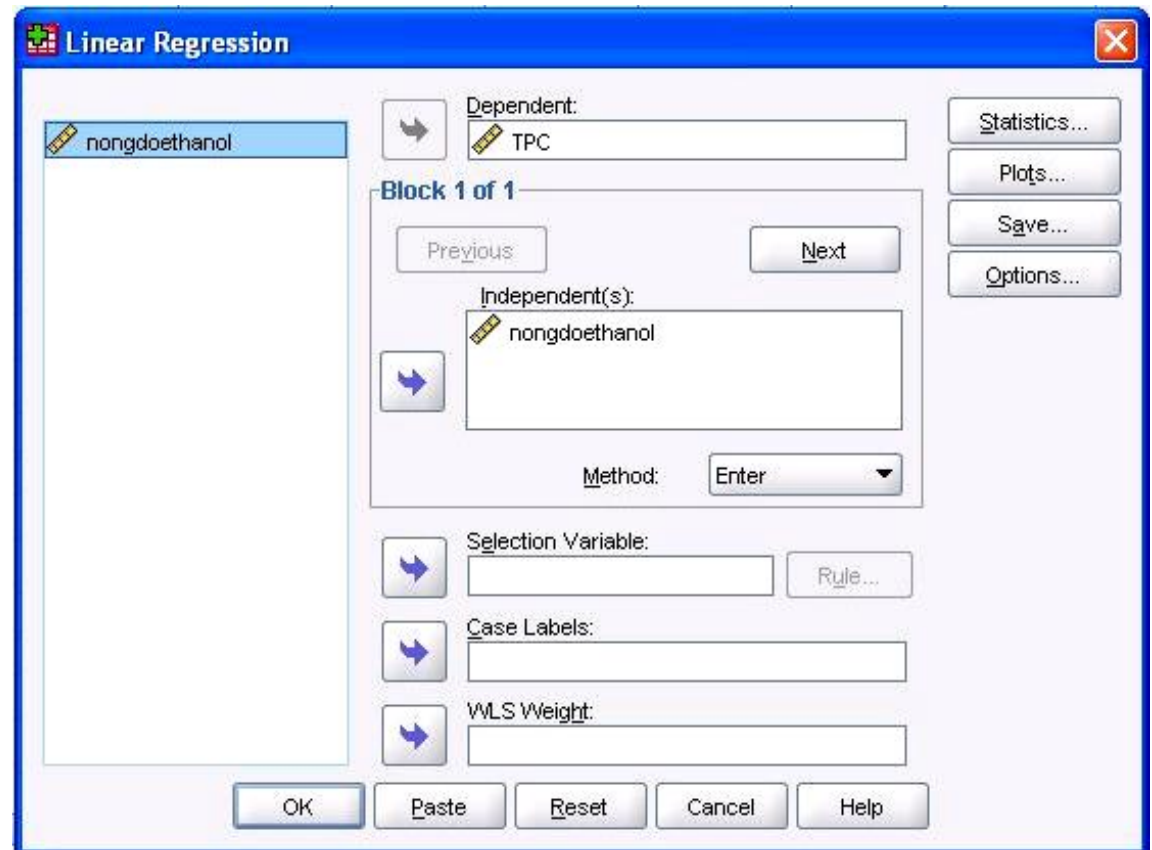
Khai báo tên biến:

+ Independent:

nongdoETHANOL

+ Dependent: TPC

(Model: Chọn linear)



## ➤ Cách đọc kết quả

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.124 <sup>a</sup>	.015	-.083	148.0426963

a. Predictors: (Constant), nongdoethanol

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3416.148	1	3416.148	.156	.701 <sup>a</sup>
	Residual	219166.399	10	21916.640		
	Total	222582.547	11			

a. Predictors: (Constant), nongdoethanol

b. Dependent Variable: TPC

## ➤ Cách đọc kết quả

**Coefficients<sup>a</sup>**

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	548.601	193.588		2.834	.018
nongdoethanol	-.995	2.520	-.124	-.395	.701

a. Dependent Variable: TPC

- Phương trình hồi quy:  $y = -0,995x + 548.610$
- Hệ số xác định:  $R^2 = 0.015$

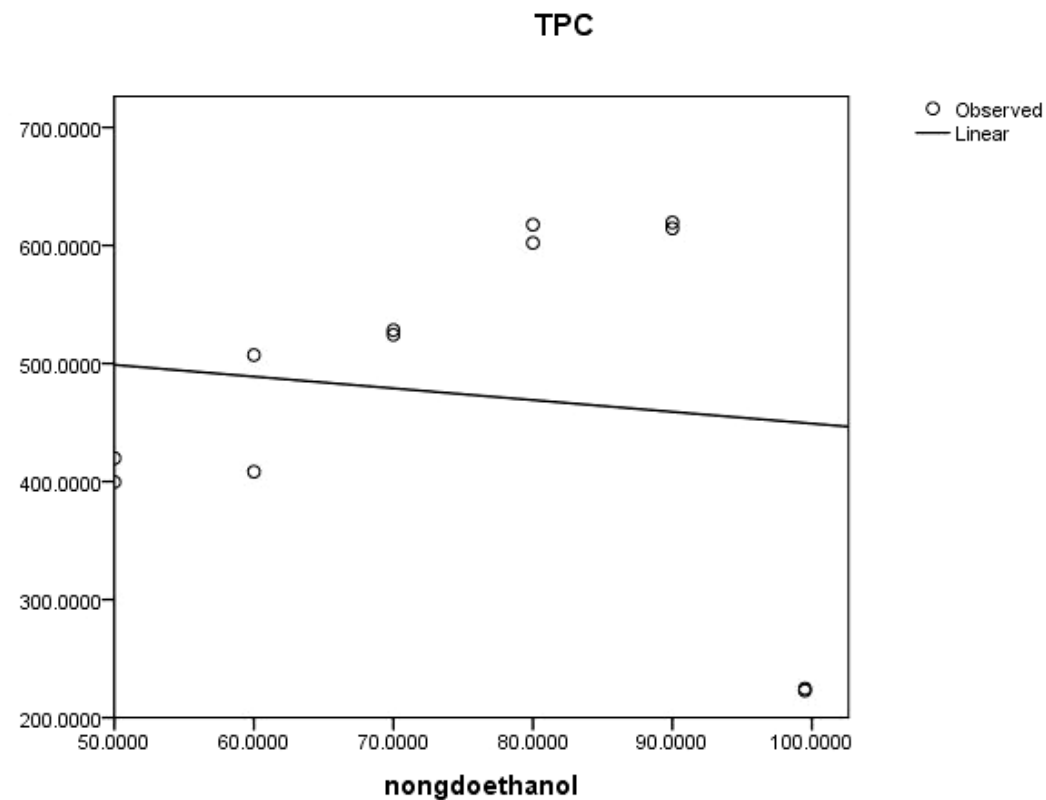
## ➤ Cách đọc kết quả

### Model Summary and Parameter Estimates

Dependent Variable: TPC

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	.015	.156	1	10	.701	548.601	-.995

The independent variable is nongdoethanol.





# Chân thành cảm ơn

