

Sử dụng bộ số liệu “HocVienCaoHoc.sav”

1. THỐNG KÊ MÔ TẢ KHOẢNG CÁCH TỪ NHÀ CỦA HỌC VIÊN ĐẾN LỚP HỌC

Descriptives			Statistic	Std. Error
khcach	Mean		4.5716	.69083
	95% Confidence Interval for Mean	Lower Bound	3.1793	
		Upper Bound	5.9638	
	5% Trimmed Mean		4.0759	
	Median		3.0000	
	Variance		21.476	
	Std. Deviation		4.63423	
	Minimum		.05	
	Maximum		25.00	
	Range		24.95	
	Interquartile Range		6.00	
	Skewness		2.142	.354
	Kurtosis		7.355	.695

Từ bảng **Descriptive** ta một số kết quả sau quan trọng sau:

- Khoảng cách trung bình 4,6716m; độ lệch chuẩn 4,63423m; khoảng tin cậy 95% cho khoảng cách trung bình là (3,1793;5,9638).
- Hệ số Kurtosis (7,355) nằm ngoài khoảng từ 1 đến 3 nên dữ liệu không xấp xỉ phân phối chuẩn.

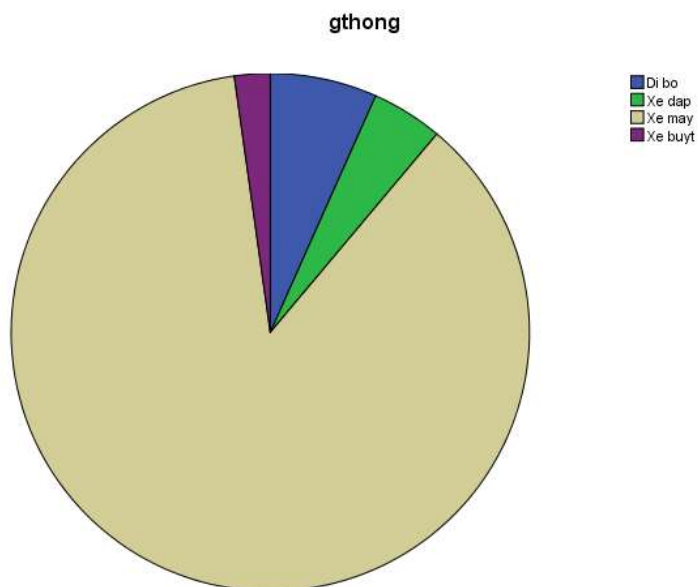
2. THỐNG KÊ MÔ TẢ PHƯƠNG TIỆN DI CHUYỂN TỪ NHÀ ĐẾN LỚP CỦA HỌC VIÊN

gthong

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Đi bộ	3	6.7	6.7	6.7
	Xe đạp	2	4.4	4.4	11.1
	Xe máy	39	86.7	86.7	97.8
	Xe buýt	1	2.2	2.2	100.0
	Total	45	100.0	100.0	

Bảng trên cho biết tần số và tần suất xuất hiện của các phương tiện di chuyển trong số liệu, cụ thể: đi bộ 6,7%; xe đạp 4,4%; xe máy 86,7%; xe buýt 2,2%. Điều đó cho thấy xe máy được sử dụng nhiều nhất (86,7%).

Biểu đồ hình tròn thể hiện tần suất của các phương tiện di chuyển:



3. SO SÁNH TỶ LỆ NAM VÀ NỮ THEO YÊU THÍCH CHƠI THỂ THAO

Giả thuyết: Tỷ lệ nữ yêu thích chơi thể thao và không yêu thích chơi thể thao là như nhau.

Đôi thuyết: Tỷ lệ nữ yêu thích chơi thể thao và không yêu thích chơi thể thao là khác nhau.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	12.857 ^a	1	.000		
Continuity Correction ^b	10.318	1	.001		
Likelihood Ratio	16.354	1	.000		
Fisher's Exact Test				.000	.000
N of Valid Cases ^b	45				

Bảng **Chi – Square Tests** cho kết quả: Kiểm định hệ số tương quan Pearson Chi – Square cho giá trị $Sig. = 0 < 0,05$ nên ta có cơ sở để cho rằng giới tính và yêu thích chơi thể thao có mối liên hệ với nhau.

thethao * gioitinh Crosstabulation

			gioitinh		Total
			nam	nu	
thethao	kh«ng thÝch	Count	12	9	21
		% within gioitinh	33.3%	100.0%	46.7%
	ThÝch	Count	24	0	24
		% within gioitinh	66.7%	.0%	53.3%
Total	Count	36	9	45	
	% within gioitinh	100.0%	100.0%	100.0%	

Bảng **thethao * gioitinh Crosstabulation** cho thấy tỷ lệ nữ nhóm không thích chơi thể thao là 100% và tỷ lệ nữ không thích chơi thể thao là 0%. Từ đó, ta có cơ sở cho rằng tỷ lệ nữ không thích chơi thể thao khác, thậm chí là cao hơn tỷ lệ nữ thích chơi thể thao.

4. SO SÁNH CHIỀU CAO TRUNG BÌNH GIỮA HỌC VIÊN THÍCH CHƠI THỂ THAO VÀ KHÔNG THÍCH CHƠI THỂ THAO

Giả thuyết: Chiều cao trung bình của học viên giữa hai nhóm thích và không thích chơi thể thao là như nhau.

Đối thuyết: Chiều cao trung bình của học viên giữa hai nhóm thích và không thích chơi thể thao là khác nhau.

Group Statistics					
thethao		N	Mean	Std. Deviation	Std. Error Mean
chieucao	kh«ng thÝch	21	161.33	6.590	1.438
	ThÝch	24	168.50	3.648	.745

Bảng **Group Statistics** cho biết các tham số thống kê cơ bản giữa các nhóm: trung bình và độ lệch chuẩn của nhóm không thích lần lượt là 161,33cm và 6,59cm; tương tự đối với nhóm thích là 168,5cm và 3,648cm.

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
									Lower Upper
chieucao	Equal variances assumed	14.155	.001	-4.589	43	.000	-7.167	1.562	-10.316 -4.017
	Equal variances not assumed			-4.425	30.267	.000	-7.167	1.619	-10.473 -3.861

Bảng **Independent Samples Test** cho biết các kết quả sau:

- Kiểm định Levene cho giá trị $Sig. = 0,001 < 0,05$ nên ta có cơ sở cho rằng phương sai của hai nhóm không bằng nhau. Do đó, ta dựa vào hàng Equal variances not assumed để kiểm tra trung bình hai nhóm.
- Kiểm định t – test cho giá trị $Sig. = 0,000 < 0,05$ nên ta có cơ sở bác bỏ giả thiết và chấp nhận đối thiết, tức là trung bình giữa hai nhóm không bằng nhau.
- Tự bảng **Group Statistics** ta thấy $mean(nhóm không thích) = 161,33 < 168,5 = mean(nhóm thích)$ nên ta có cơ sở cho rằng chiều cao trung bình của học viên thích chơi thể thao cao hơn chiều cao trung bình của học viên không thích chơi thể thao.

5. PHÂN TÍCH HỒI QUY TUYẾN TÍNH BIẾN THỜI GIAN TỪ NHÀ ĐẾN LỚP THEO CÁC BIẾN CHIỀU CAO, CÂN NẶNG, PHƯƠNG TIỆN GIAO THÔNG VÀ KHOẢNG CÁCH.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.818 ^a	.669	.636	6.23035	1.342

a. Predictors: (Constant), khcach, chieucao, gthong, kannang

b. Dependent Variable: thgian

Bảng **Model Summary** cho biết các kết quả sau:

- Hệ số tương quan R = 0,818 khá cao cho thấy có mối liên hệ tuyến tính chặt chẽ giữa biến phụ thuộc với các biến độc lập.
- Hệ số Adjusted R Square = 0,636 cho biết mô hình giải thích được 63,6% sự biến thiên của dữ liệu

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3141.310	4	785.328	20.231	.000 ^a
	Residual	1552.690	40	38.817		
	Total	4694.000	44			

Bảng **ANOVA** cho giá trị Sig. = 0,000 < 0,05 nên ta có cơ sở cho rằng các hệ số hồi quy là khác 0, tức là tồn tại mối tương quan tuyến tính giữa biến phụ thuộc với các biến độc lập.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.227	28.209		.114	.909
	chieucao	.123	.213	.075	.578	.567
	kannang	-.278	.185	-.193	-1.506	.140
	gthong	-.562	1.816	-.031	-.309	.759
	khcach	1.895	.217	.850	8.747	.000

Bảng **Coefficients** cho biết các kết quả sau:

- Chỉ có giá trị Sig. của biến khoảng cách là nhỏ hơn 0,05 (0,000) nên chỉ có biến này có ý nghĩa trong mô hình. Các biến còn lại không có ý nghĩa.
- Mô hình ước lượng được:

$$\text{Thời gian} = 3,227 + 0,123 * (\text{Chiều cao}) - 0,278 * (\text{cân nặng}) - 0,562 * (\text{phương tiện}) + 1,895 * (\text{khoảng cách})$$

- Từ mô hình trên, ta có nhận xét như sau: Trong điều kiện các yếu tố khác không đổi, nếu tăng khoảng cách thêm 1km thì thời gian sẽ tăng thêm 1,895 phút.