



ĐẠI HỌC FPT

COMPUTER PROJECT

MAS192 - SE1622

***Đề tài : Khảo sát chiều cao, cân nặng của
sinh viên Nam và Nữ Đại Học FPT***

Giảng Viên : Nguyễn Việt Anh

Thành viên nhóm :

- 1. Đào Xuân Phúc : HE153725**
- 2. Bùi Thái Dương : HE153717**
- 3. Nguyễn Văn Quyết : HE153773**

Bản báo cáo project !

A. Phần mở đầu project

Project được thực hiện bởi nhóm dưới sự dẫn dắt của giảng viên Nguyễn Việt Anh với mục tiêu áp dụng các kiến thức được học trong bộ môn MAS291 để nghiên cứu về đề tài mà cả nhóm quan tâm tới. Phạm vi kiến thức sử dụng trong project bao gồm có các chương 8, 9, 10 và chương 11 .

B. Báo cáo quá trình thực hiện project

1. Lý do chọn đề tài

Sau quá trình thảo luận, cùng nhau đóng góp ý kiến, nhóm em muốn làm một chủ đề thực tế , có liên hệ trực tiếp tới cuộc sống hàng ngày. Nhóm em muốn khảo sát về chiều cao, cân nặng của sinh viên trong Đại Học FPT để giúp mọi người giải đáp thắc mắc về vấn đề đó. Vấn đề về chiều cao cân nặng luôn được mọi người quan tâm chú ý tới không chỉ các bạn nữ mà cả các bạn nam . Đó hơn hết là vẻ bề ngoài của mọi người , điều được các bạn sinh viên rất dễ tâm tới . Hơn thế nữa chiều cao , cân nặng cũng là một khái niệm thước đo trong việc khám sức khỏe của mọi người . Nhóm sẽ giúp mọi người hiểu hơn về :

- Chiều cao, cân nặng trung bình của các bạn Nam và Nữ trong Đại Học FPT.
- Đánh giá , so sánh chiều cao, cân nặng của Nam và Nữ trong Đại Học FPT .
- Tính khoảng tin cậy cho giá trị trung bình chiều cao, cân nặng của Nam và Nữ trong Đại Học FPT.
- Mối tương quan giữa chiều cao và cân nặng của sinh viên FPT.

2. *Nguồn Data*

Để tìm được nguồn dữ liệu đáng tin cậy, chính xác và thực tế nhất. Nhóm em đã làm một google form đặt ra những câu hỏi liên quan tới đề tài và nhờ các bạn Nam và Nữ có thể điền vào đó .

- Link khảo sát:

https://docs.google.com/forms/d/14fnmhcVsMDULGx2Hlc6fVkrCqePbFpJHhfWbgM2aNzA/edit?usp=forms_home&ths=true

- Đường link kết quả khảo sát :

https://docs.google.com/spreadsheets/d/1CQ1Sfqx5mNGReZJB_ROfzRJvrf8xw3-tYdUm5Bqbw/edit?resourcekey#gid=1629398007

- Độ tin cậy :

- Đây là quá trình khảo sát thực tế tới các bạn nên nó rất chính xác .

- Đây là một dữ liệu nhóm tự khảo sát thực tế nên chất lượng thực tế là rất cao, được chọn lọc kỹ càng , sát với ngoại hình của mọi người nhất.

3. Tiến hành các mục trong project

****. Xác định dân số và mẫu:***

- Dân số ở đây là tất cả sinh viên ĐH FPT
- Mẫu ở đây là 36 sinh viên được khảo sát thông qua Google Form

Câu 1:

1. Xây dựng khoảng tin cậy 95% cho trung bình chiều cao của sinh viên ĐH FPT

Áp dụng công thức:

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\Rightarrow 165.58 - z_{0,025} * \frac{8.16}{\sqrt{36}} \leq \mu \leq 165.58 + z_{0,025} * \frac{8.16}{\sqrt{36}}$$

$$\Rightarrow 162.91 \leq \mu \leq 168.25$$

2. Xây dựng khoảng tin cậy 95% cho trung bình cân nặng của sinh viên ĐH FPT

Áp dụng công thức:

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\Rightarrow 57.81 - z_{0.025} * \frac{10.37}{\sqrt{36}} \leq \mu \leq 57.81 + z_{0.025} * \frac{10.37}{\sqrt{36}}$$

$$\Rightarrow 54.42 \leq \mu \leq 61.2$$

Câu 2: Có phải trung bình chiều cao của nam giới trong trường ĐH FPT lớn hơn 8cm so với trung bình chiều cao của nữ giới trong trường ĐH FPT hay không với $\alpha = 0.05$

Ta có :

$$H_0: \mu_1 - \mu_2 = \Delta_0 \text{ với } \Delta_0 = 8$$

$$H_1: \mu_1 - \mu_2 > \Delta_0$$

$$\bar{x}_1 = 171.5 \text{ cm}, s_1 = 4.958, n_1 = 18$$

$$\bar{x}_2 = 160.556 \text{ cm}, s_2 = 6.68, n_2 = 18$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$\Rightarrow S_p = 5.88$$

Test statistic được tính theo công thức:

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Sử dụng công thức có: $T_0 = 1.5$

Ta có:

Vì $T_0 = 1.5 < T_{0,05,34} = 1.691$, nên ta sẽ chấp nhận H_0 . Vậy ta không có đủ chứng cứ để kết luận rằng trung bình chiều cao của nam giới ĐH FPT cao hơn 8 cm so với trung bình chiều cao của nữ giới ĐH FPT với $\alpha = 0.05$

Câu 3: Có phải trung bình cân nặng của nam giới trong trường ĐH FPT lớn hơn 5kg so với trung bình cân nặng của nữ giới trong trường ĐH FPT hay không với $\alpha = 0.05$

Ta có :

$$H_0: \mu_1 - \mu_2 = \Delta_0 \text{ với } \Delta_0 = 5$$

$$H_1: \mu_1 - \mu_2 > \Delta_0$$

$$X_1 = 63.83 \text{ kg}, s_1 = 9.92, n_1 = 18$$

$$X_2 = 51.11, s_2 = 7.28, n_2 = 18$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$\rightarrow S_p = 8.7$$

Test statistic được tính theo công thức:

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Sử dụng công thức có: $T_0 = 2.66$

Ta có:

Vì $T_0 = 2.66 > T_{0,05,34} = 1.691$, nên ta sẽ bác bỏ H_0 . Vậy ta có đủ chứng cứ để kết luận rằng trung bình cân nặng của nam giới ĐH FPT cao hơn 5kg so với trung bình cân nặng của nữ giới ĐH FPT với $\alpha = 0.05$

Câu 4: Có phải trung bình cân nặng của sinh viên FPT có lớn hơn 55kg hay không với $\alpha = 0.05$

Ta có :

$$H_0: = 55$$

$$H_1 : > 55$$

$$\bar{x} = 57.81 \text{ kg}, \mu = \mu_0 = 55 \text{ kg}, s = 10.37, n = 36$$

Test statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$\Rightarrow T_0 = 1.63$$

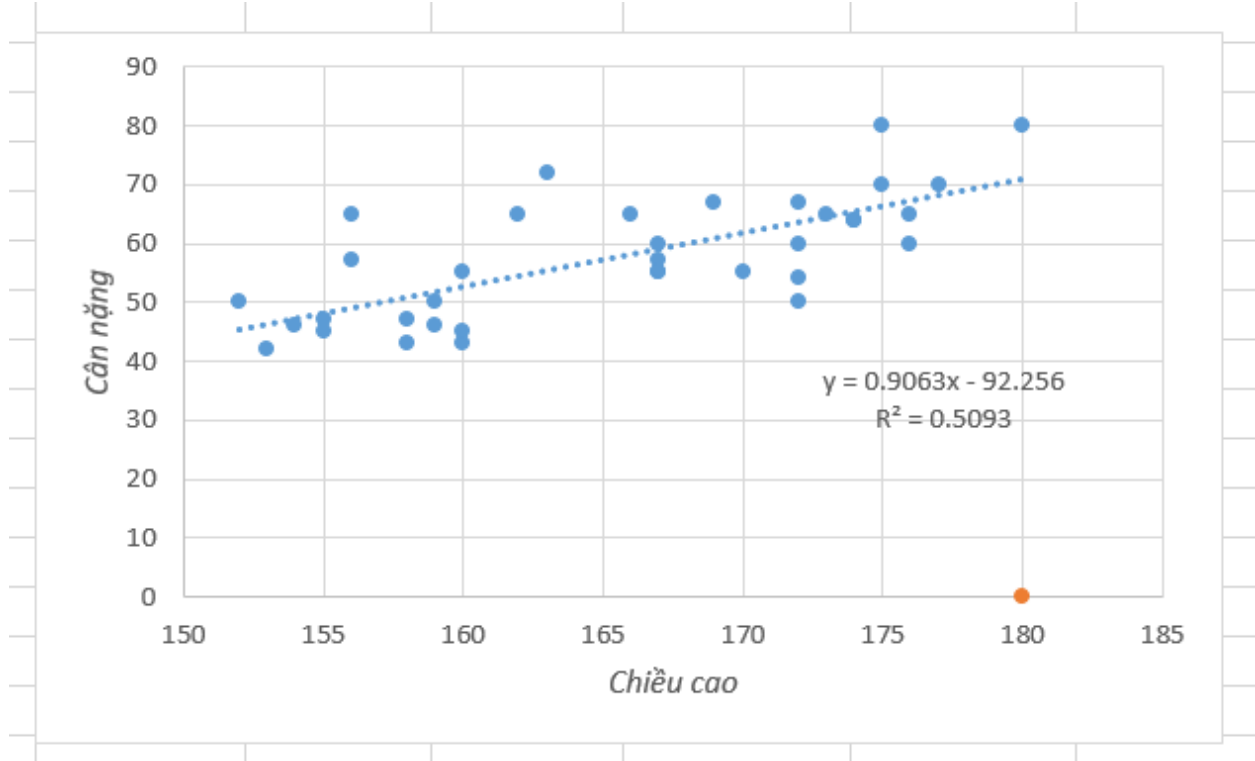
$$\text{Ta có: } T_{0,05,35} = 1.69$$

Vì $1.63 < 1.69$ nên ta sẽ chấp nhận H_0 . Vậy chưa có đủ bằng chứng để kết luận rằng trung bình cân nặng của sinh viên FPT lớn hơn 55kg

Hỏi quy tuyến tính:

Câu 5: Liệu có sự tương quan nào giữa chiều cao và cân nặng của sinh viên FPT hay không?

- Ta có Scatter plot:



- Giả sử mô hình tuyến tính là hợp lý. Ta có phương trình tuyến tính đơn giản: $y = \hat{\beta}_0 + \hat{\beta}_1 x$
- Từ số liệu thu thập được với Y đại diện cho cân nặng và X đại diện cho chiều cao ta có ước tính bình phương nhỏ nhất của các hệ số trong mô hình hồi quy tuyến tính đơn giản:

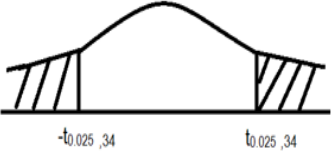
$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = 0.906$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 = -92.206$$

- Áp dụng kiểm định giả thiết để kiểm tra mức độ hợp lý của mô hình tuyến tính đơn giản:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

Regression Statistics									
Multiple R	0.713673305								
R Square	0.509329587								
Adjusted R Square	0.494898104								
Standard Error	7.367905176								
Observations	36								
ANOVA									
	df	SS	MS	F	Significance F				
Regression	1	1915.913981	1915.913981	35.292949	1.0279E-06				
Residual	34	1845.724907	54.28602669						
Total	35	3761.638889							
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
Intercept	-92.25637362	25.28943604	-3.6480202	0.0008767	-143.65069	-40.86205608	-143.6506912	-40.86205608	
Chiều cao	0.90626228	0.152549198	5.940786916	1.028E-06	0.59624501	1.21627955	0.59624501	1.21627955	

Sử dụng biến t-Tests với độ tin cậy $\alpha = 0.05$:

$$t_{\frac{\alpha}{2}, n-2} = 2.032$$

$$SS_e = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 1845.7249$$

$$MS_e = \sigma^2 = \frac{SS_e}{n-2} = 54.286$$

$$S_{xx} = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} = 2332.75$$

$$t_0 = \frac{\hat{\beta}_1 - 0}{\sqrt{\frac{\sigma^2}{S_{xx}}}} = 5.94$$

H_0 bị bác bỏ vì: $|t_0| > t_{\frac{\alpha}{2}, n-2}$

Việc bác bỏ H_0 dẫn tới có mối quan hệ tuyến tính giữa chiều cao và cân nặng. Vì vậy từ số liệu, ta có thể ước tính mức độ tương quan giữa chiều cao và cân nặng của sinh viên FPT theo phương trình tuyến tính:

$$y = 0.9063x - 92.256$$

C. Đánh giá kết quả nghiên cứu

Sau khi khảo sát thực tế, nghiên cứu và tính toán theo các kiến thức đã học trong môn MAS291, nhóm chúng em đã đưa ra kết luận như sau :

- Chênh lệch trung bình chiều cao và cân nặng giữa nam và nữ trường ĐH FPT chênh lệch nhau không quá nhiều
- Trung bình cân nặng của sinh viên FPT không quá cao