

Phân tích xu hướng và các yếu tố tác động đến việc nộp chứng chỉ tiếng Anh tại trường Đại học

Analysis of the current situation of English Certificate Submission among Students

Trần Thành Trung, Cao Phi Phụng

Khoa Công nghệ thông tin, Trường Đại học Tôn Đức Thắng

Tóm tắt

Chứng chỉ tiếng Anh quốc tế đáp ứng nhu cầu đầu ra của chương trình học là vấn đề được nhiều trường đại học quan tâm để đảm bảo chất lượng của sinh viên trong bối cảnh hội nhập quốc tế. Bài báo cáo phân tích xu hướng nộp tiếng Anh của sinh viên từ năm 2016 đến năm 2023, xem xét các yếu tố về giới tính, khu vực, nơi ở, xuất thân gia đình có ảnh hưởng đến việc nộp chứng chỉ tiếng Anh cũng như trả lời các câu hỏi về việc học tập tại các thành phố lớn hay các tỉnh thành khác có tác động đến việc nộp tiếng Anh hay không?

Abstract

International English certificates that meet the output needs of the curriculum are an issue that many universities are concerned about to ensure the quality of students in the context of international integration. The report analyzes the trends in students's English submission from 2016 to 2023, considering factors such as gender, region, place of residence, and family background that affect the submission of English certificates as well as such as answering questions about whether studying in big cities or other provinces affects English submission?.

Từ khoá: chứng chỉ tiếng Anh, đại học, sinh viên.

Keywords: chứng chỉ tiếng Anh, sinh viên, đại học.

1. GIỚI THIỆU

Đề tài này tập trung phân tích dữ liệu thống kê sinh viên nộp chứng chỉ tiếng Anh đầu ra của một trường đại học trên địa bàn Thành phố Hồ Chí Minh. Bộ dữ liệu này được thu thập từ 38.000 sinh viên thuộc các khóa tuyển sinh từ năm 2016 đến năm 2023. Thông tin này bao gồm: *giới tính sinh viên, khu vực sinh sống, nghề nghiệp phụ huynh và loại chứng chỉ tiếng Anh và thời điểm nộp chứng chỉ*. Từ dữ liệu này, chúng tôi sẽ tiến hành phân tích, nghiên cứu các yếu tố tác động đến việc nộp chứng chỉ tiếng Anh đầu ra của SV. Chúng tôi đã đặt ra các câu hỏi nghiên cứu cụ thể là:

- Việc sinh viên sinh sống và học tập tại các thành phố lớn, có được nhiều điều kiện học tiếng Anh nên việc nộp chứng chỉ tiếng Anh của các sinh viên sinh sống tại đây có sự khác

biệt so với các sinh viên sinh sống ở các tỉnh thành khác không?

- Sinh viên nữ có khả năng học ngoại ngữ tốt hơn và nộp chứng chỉ tiếng Anh đầu ra sớm hơn sinh viên nam không?
- Nghề nghiệp của cha (được xem là lao động chính trong gia đình) thuộc nhóm ngành kinh doanh, buôn bán có phải là yếu tố tác động đến điều kiện tài chính giúp sinh viên có thể học tập, ôn luyện để nộp chứng chỉ tiếng Anh đúng hạn hay không?
- Xu hướng lựa chọn loại chứng chỉ tiếng Anh hiện nay của sinh viên có thay đổi như thế nào từ năm 2016 đến năm 2023?

Đây là các câu hỏi nghiên cứu cần được làm rõ trong báo cáo này.

2. PHÂN TÍCH DỮ LIỆU

2.1. Mô tả dữ liệu

Thông tin tập dữ liệu thông tin sinh viên nộp chứng chỉ tiếng Anh như sau:

- Nguồn dữ liệu: Dữ liệu thứ cấp được thu thập từ các sinh viên thuộc khóa tuyển sinh từ năm 2016 đến năm 2023 từ một trường Đại học trên địa bàn Thành phố Hồ Chí Minh.
- Số lượng: gần 38.000 dòng.
Tập dữ liệu gồm các cột (biến) như sau:
 - **Giới tính (gioi tinh):** là kiểu dữ liệu định tính gồm Nam, Nữ.
 - **Ngành (nganh):** là kiểu dữ liệu định tính gồm 36 ngành đào tạo khác nhau.
 - **Năm TS (nam_ts):** là kiểu dữ liệu định lượng rời rạc có các năm 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, thể hiện năm mà sinh viên được tuyển sinh vào trường.
 - **Khu vực (khu_vuc):** là kiểu dữ liệu định tính thứ tự gồm: 1, 2, 2NT, 3 thể hiện khu vực tuyển sinh đại học tại Việt Nam.
 - **Nơi học THPT (noi_hoc_thpt):** là kiểu dữ liệu định tính thể hiện nơi tỉnh/thành phố sinh viên học Trung học phổ thông.
 - **Phân loại nghề nghiệp cha (phan_loai_nghe_nghiep_cha):** là kiểu định tính thể hiện ngành nghề làm việc của cha.
 - **Phân loại nghề nghiệp mẹ (phan_loai_nghe_nghiep_me):** tương tự như Phân loại nghề nghiệp cha, ở đây dữ liệu thể hiện ngành nghề làm việc của mẹ.
 - **Tiếng Anh đầu ra:** là kiểu dữ liệu định tính, khi sinh viên đã nộp chứng chỉ tiếng Anh đầu ra, dữ liệu có cấu trúc: Số: [số thông báo]/TB, ngày [ngày công nhận] – [loại chứng chỉ]:[điểm tổng]. Nếu để trống là sinh viên chưa nộp chứng chỉ tiếng Anh đầu ra.

2.2. Tiền xử lý dữ liệu

Thực hiện chuẩn hóa các thuộc tính:

- Giới tính: Nam là 0; Nữ là 1;
- Ngành: mã hóa bằng mô hình Label Encoder.
- Khu vực: chuyển đổi khu vực “2NT” thành 4.

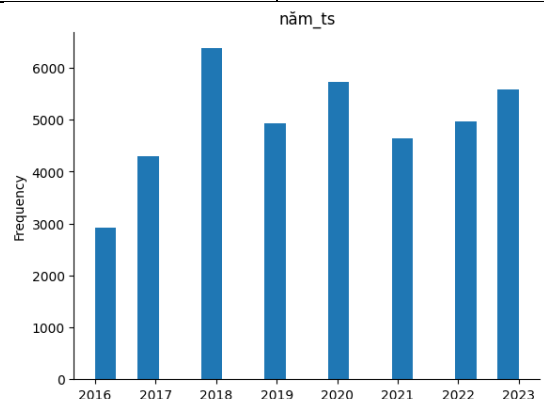
- Thực hiện tách thông tin từ cột dữ liệu “**Tiếng Anh đầu ra**”. Sử dụng chức năng FlashFill của Excel tách các trường thông tin: [ngày công nhận], [loại chứng chỉ], [điểm chứng chỉ].
- Tạo thêm cột **Thời điểm nộp chứng chỉ tiếng Anh (nam_ta_dau_ra)** theo công thức: Thời điểm nộp chứng chỉ tiếng Anh = Năm tuyển sinh - Năm nộp chứng chỉ tiếng Anh đầu ra.

2.3. Phân tích thống kê mô tả

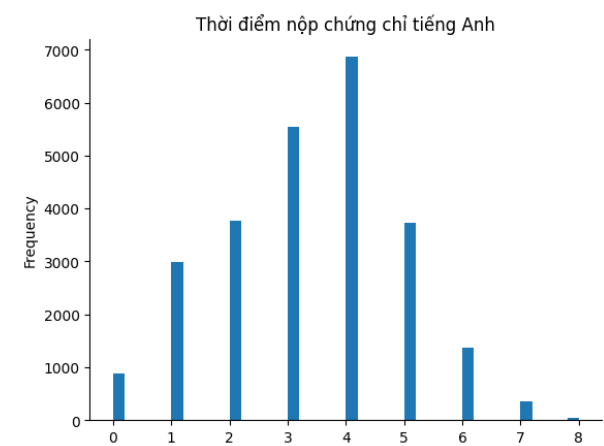
Để phân tích tổng quan dữ liệu chúng tôi đã thu thập được, chúng tôi trình bày một số thống kê sơ lược và đưa ra nhận xét liên quan đến dữ liệu như sau:

Thống kê số lượng sinh viên qua các năm:

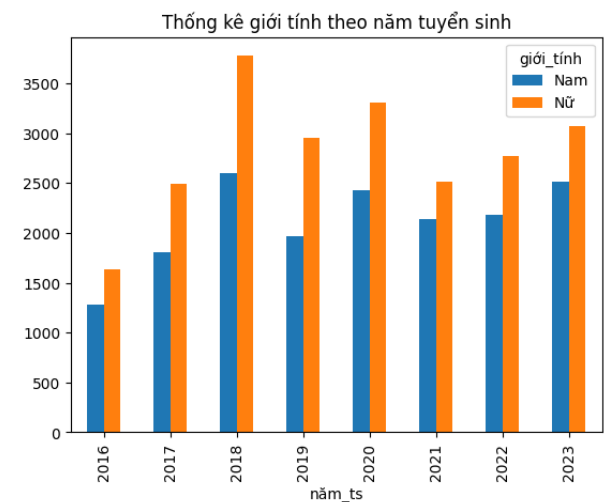
Năm	Số lượng sinh viên
2016	2916
2017	4294
2018	6376
2019	4923
2020	5737
2021	4651
2022	4960
2023	5591



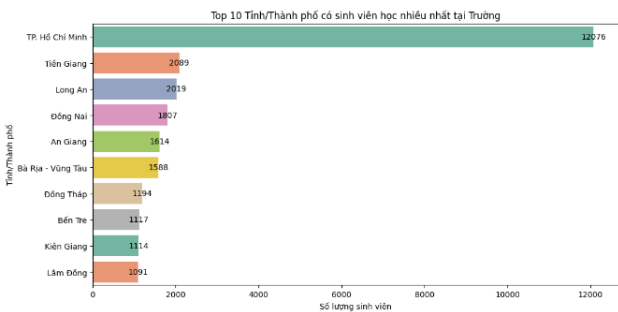
Biểu đồ thống kê số lượng SV theo các năm



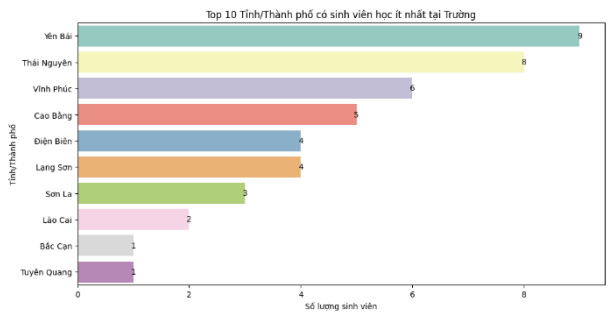
Biểu đồ số lượng nộp chứng chỉ tiếng Anh tại các thời điểm



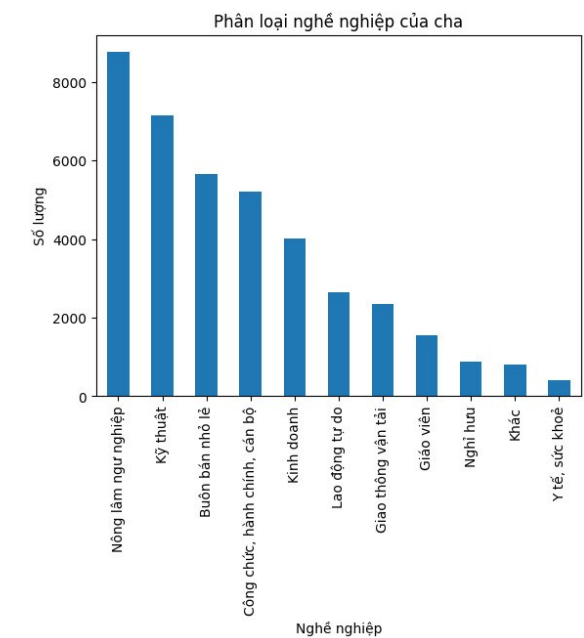
Biểu đồ thống kê số lượng sinh viên theo giới tính qua các năm



Biểu đồ thống kê TOP 10 Tỉnh/Thành phố có số lượng sinh viên theo học nhiều nhất



Biểu đồ TOP 10 Tỉnh/Thành phố có số lượng sinh viên theo học ít nhất



Biểu đồ thống kê số số lượng SV theo ngành nghề của cha

Sau đây là bảng thống kê mô tả dữ liệu thông tin năm nộp chứng chỉ tiếng Anh và điểm của 03 loại chứng chỉ thông dụng nhất.

	Năm nộp CCTA	Điểm APTIS	Điểm TOEIC	Điểm IELTS
Mean	3.30	151.90	619.88	5.94
Std	1.56	16.14	96.34	0.56
Min	0	104	155	5
25%	2	140	550	5.5
50%	3	155	595	6
75%	4	164	665	6
Max	8	191	995	8.5

Từ các mô tả thống kê trên, chúng tôi có nhận xét như sau:

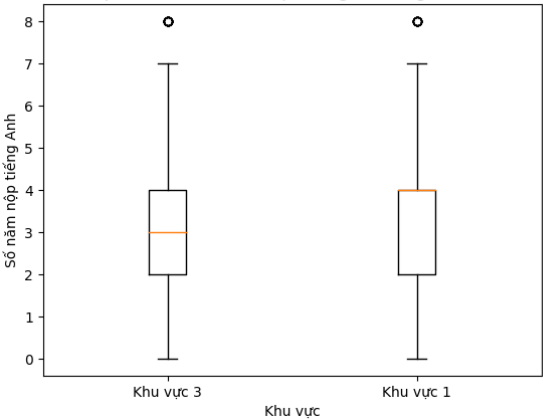
- Số lượng nộp chứng chỉ tiếng Anh qua các thời điểm thể hiện phân phối chuẩn lệch trái. Phần đông sinh viên hoàn thành chứng chỉ tiếng Anh vào năm cuối (thời điểm chuẩn bị tốt nghiệp) và dữ liệu lệch trái do sinh viên phần đông hoàn thành đúng tiến độ chương trình đào tạo 4 năm.
- Số lượng sinh viên theo học đa số từ các tỉnh/thành phố miền Nam và thưa thớt dần ở các tỉnh/thành phố miền Trung và miền Bắc.

Chúng tôi đã thực hiện kiểm định t-test để có cơ sở kết luận các giả thiết mà chúng tôi đã đặt ra ban đầu. Tổng hợp kết quả kiểm định các giả thuyết nghiên cứu được thể hiện ở bảng sau:

Giả thuyết	Kết quả	
	p_value	Kết luận
H_1 : Việc sinh viên sinh sống và học tập tại các thành phố lớn, có được nhiều điều kiện học tiếng Anh nên việc nộp chứng chỉ tiếng Anh của các sinh viên sinh sống tại đây có sự khác biệt so với các sinh viên sinh sống ở các tỉnh thành khác không?	$1.23e^{-6}$	Chấp nhận giả thuyết
H_1 : Sinh viên nữ có khả năng học ngoại ngữ tốt hơn và nộp chứng chỉ tiếng Anh đầu ra sớm hơn sinh viên nam không?	0.0005	Chấp nhận giả thuyết
H_1 : Nghề nghiệp của cha (được xem là lao động chính trong gia đình) thuộc nhóm ngành kinh doanh, buôn bán có phải là yếu tố tác động đến điều kiện tài chính giúp sinh viên có thể học tập, ôn luyện để nộp chứng chỉ tiếng Anh đúng hạn hay không?	0.014	Chấp nhận giả thuyết

2.4. Trực quan hóa dữ liệu

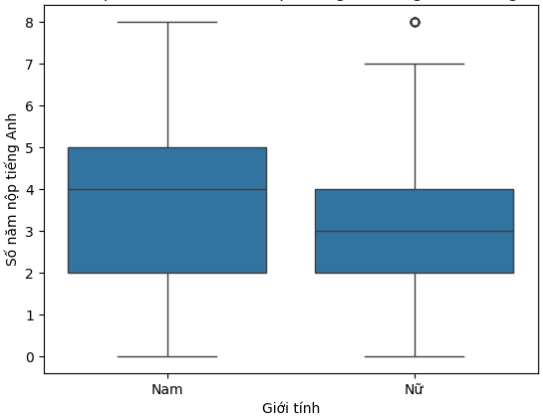
Biểu đồ boxplot của thời điểm nộp chứng chỉ tiếng Anh theo khu vực



Biểu đồ boxplot của thời điểm nộp chứng chỉ tiếng Anh theo khu vực

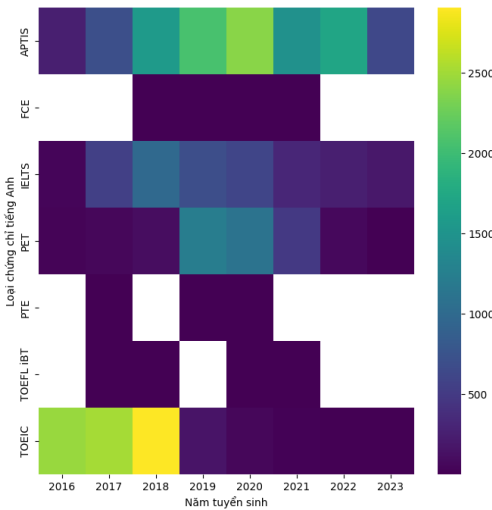
Nhìn vào biểu đồ chúng tôi có nhận xét: thời điểm nộp chứng chỉ tiếng Anh của khu vực 3 (sinh sống tại tỉnh/thành phố lớn) tập trung vào năm 3 và khu vực 1 (sinh sống tại các vùng quê) là năm 4.

Biểu đồ boxplot của thời điểm nộp chứng chỉ tiếng Anh theo giới tính



Biểu đồ boxplot của thời điểm nộp chứng chỉ tiếng Anh theo giới tính

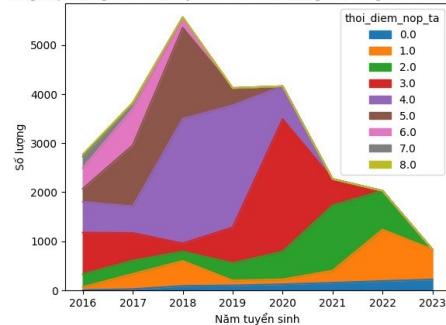
Theo biểu đồ trên, chúng tôi nhận thấy thời điểm sinh viên nữ nộp chứng chỉ tiếng Anh vào khoảng năm 2 đến năm 4 và chủ yếu tập trung vào thời điểm năm 3, còn sinh viên nam nộp chứng chỉ tiếng Anh vào phân bố từ năm 2 đến năm 5, và tập trung nộp vào thời điểm năm 4. Điều này tăng cơ sở kết luận được giả thiết sinh viên nữ có khả năng ngoại ngữ và nộp chứng chỉ tiếng Anh sớm hơn sinh viên nam.



Biểu đồ thống kê số lượng các loại chứng chỉ tiếng Anh của các khóa tuyển sinh

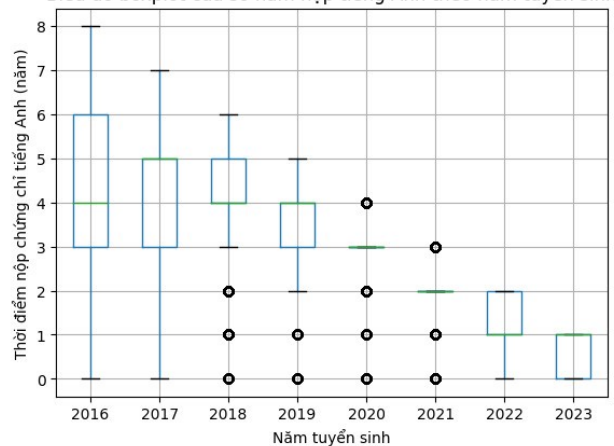
- Đối với chứng chỉ APTIS: đây là chứng chỉ tiếng Anh được công nhận vào năm 2022. Chứng chỉ được đánh giá có cấu trúc bài thi gần gũi, dễ đạt mức yêu cầu. Vì vậy, các khóa tuyển sinh năm 2020 (sinh viên năm 3 vào thời điểm 2022) có xu hướng chọn để thi hơn so với các chứng chỉ khác.
- Đối với chứng chỉ IELTS: đây luôn được sinh viên lựa chọn trong khoảng trên dưới 1500 sinh viên mỗi khóa. Vì đây là chứng chỉ uy tín thông dụng, được nhiều quốc gia công nhận, lợi thế sử dụng cho nhiều mục đích khác nhau như xét học bổng, đi làm, định cư, du học...

Biểu đồ vùng xếp chồng của năm tuyển sinh với số lượng của từng thời điểm nộp chứng chỉ TA



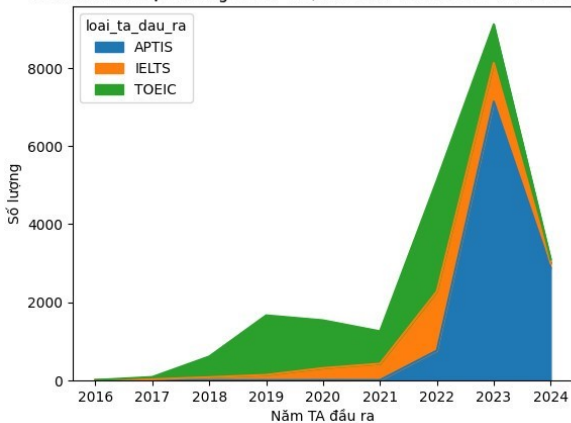
Biểu đồ vùng xếp chồng của khóa tuyển sinh với số lượng của từng thời điểm nộp chứng chỉ

Boxplot grouped by nam ts
Biểu đồ boxplot của số năm nộp tiếng Anh theo năm tuyển sinh



Biểu đồ boxplot thể hiện thời điểm nộp tiếng Anh của của từng khóa tuyển sinh

Biểu đồ thể hiện chứng chỉ APTIS, IELTS và TOEIC theo thời điểm nộp

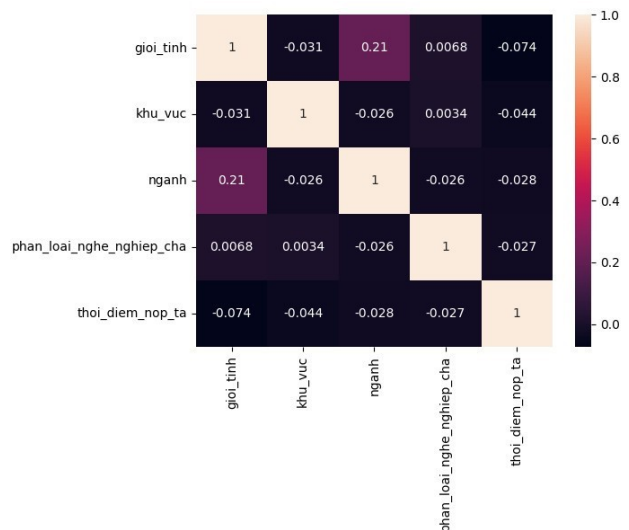


Phân bố số lượng nộp các chứng chỉ tiếng Anh qua các năm

Thông qua 02 biểu đồ trên, chúng tôi nhận định xu hướng thi chứng chỉ tiếng Anh của sinh viên các khóa như sau:

- Đối với chứng chỉ TOEIC: được ưa chuộng với khóa tuyển sinh năm 2016, tuy nhiên số lượng nộp chứng chỉ TOEIC giảm dần qua các khóa, nguyên nhân do sự thay đổi quy định công nhận chứng chỉ TOEIC từ thi đạt 2 kỹ năng (nghe, đọc) sang 4 kỹ năng (nghe nói đọc viết) khiến việc thi đạt chứng chỉ khó hơn.

Đối với hai biểu đồ trên, chúng tôi nhận xét rằng thời điểm nộp chứng chỉ tiếng Anh có xu hướng sớm dần từ khóa tuyển sinh 2020. Điều này cho thấy ý thức của sinh viên nộp chứng chỉ tiếng Anh sớm qua các khóa. Bên cạnh đó, từ năm 2022 khi chứng chỉ tiếng Anh APTIS được đưa vào áp dụng thì đa số SV có xu hướng lựa chọn nộp chứng chỉ APTIS, đây được xem là một trong các yếu tố hỗ trợ sinh viên nộp chứng chỉ tiếng Anh đúng hạn.



Biểu đồ thể hiện hệ số tương quan giữa các biến với biến phụ thuộc là thời điểm nộp chứng chỉ Tiếng Anh

	coef	std err
const	3.7370	0.059
gioi_tinh	-0.2313	0.037
khu_vuc	-0.0654	0.016
ngành	-0.0023	0.002
phan_loai_nghe_nghiep_cha	-0.0175	0.007

Kết quả chạy mô hình hồi quy tuyến tính giữa các biến độc lập và biến phụ thuộc (thời điểm nộp chứng chỉ tiếng Anh)

Với biểu đồ hệ số tương quan và kết quả chạy mô hình trên, chúng tôi nhận xét rằng: đối với biến phụ thuộc là ‘Thời điểm nộp tiếng Anh’, biến ‘Giới tính’ có mức độ tác động cao nhất.

3. KẾT LUẬN

Từ các phân tích phía trên, chúng tôi kết luận rằng thời điểm sinh viên nộp chứng chỉ tiếng Anh có sự ảnh hưởng từ các yếu tố như giới tính của sinh viên, khu vực sinh sống và nghề nghiệp của cha. Yếu tố giới tính có tác động mạnh mẽ nhất đến thời điểm sinh viên nộp tiếng Anh. Ngoài ra, xu hướng lựa chọn chứng chỉ tiếng Anh cũng làm thay đổi thời điểm nộp chứng chỉ tiếng Anh của sinh viên.