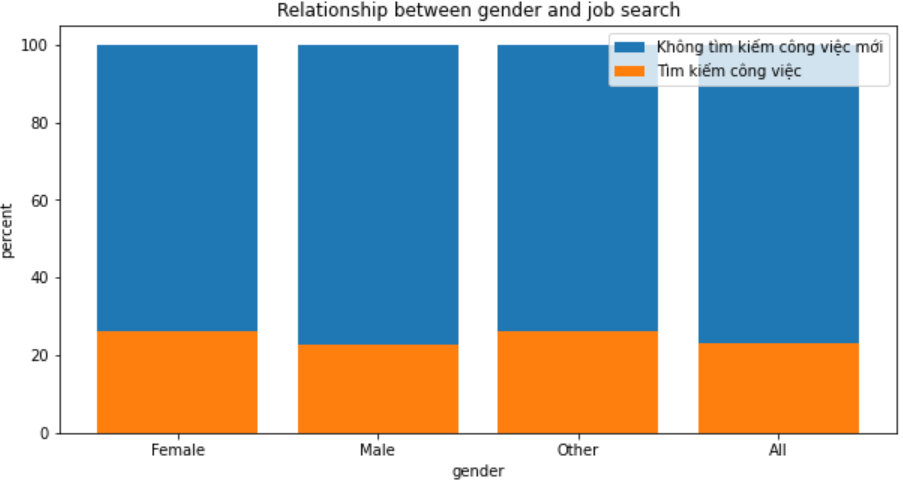
1. **Thực nghiệm, kết quả và thảo luận**
   1. **Giới tính có ảnh hưởng đến việc lựa chọn thay đổi công việc hay không?**

*Biểu đồ cột ghép thể hiện tỉ lệ quyết định lựa chọn thay đổi công việc của các nhóm giới tính*



*Nhận xét:*

* Tỉ lệ tìm kiếm công việc mới giữa các nhóm giới tính giống nhau.
* Nếu sự phân bố của những người tìm việc là như nhau trong tất cả các loại, điều đó chỉ ra rằng hai biến này độc lập với nhau.
* Ta sẽ kiểm định giả thuyết: Giới tính không ảnh hưởng đến việc thay đổi tỉ lệ tìm việc. Mức độ tin cậy 95%

*Kiểm định giả thuyết*

* Sử dụng hàm chi2\_consequency của thư việc scipy.stats, ta thu được kết quả

x^2: 9.0421817403733 p\_val: 0.010877151633874342 df: 2

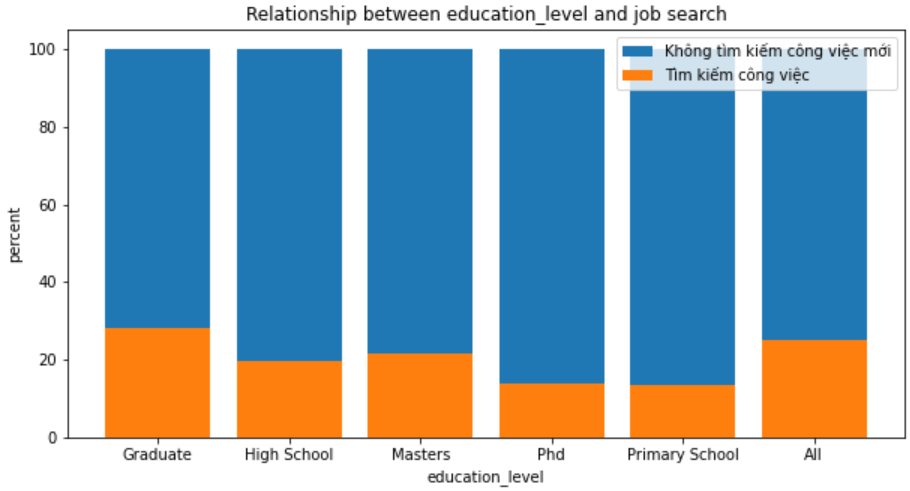
* Kết luận

+ Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Tức là, biết giới tính của ứng viên cho chúng ta thông tin về việc liệu người đó sẽ tìm kiếm một công việc mới hay không.

+ Giá trị p\_value không quá xa so với alpha. Có nghĩa là, mối quan hệ giữa hai biến không mạnh lắm. Mối quan hệ giữa hai người gần với sự độc lập chứ không phải là một mối liên kết rất bền chặt.

**6.2. Trình độ học vấn của ứng viên có ảnh hưởng đến việc thay đổi lựa chọn công việc không?**

*Biểu đồ cột ghép thể hiện mối tương quan giữa biến education\_level và target*



*Nhận xét*

* Tỉ lệ người tìm việc mới gần như giống nhau ở tất cả các cấp học
* Sự phân bố của những người tìm viêc ở các nhóm trình độ học vấn khác nhau dường như là như nhau
* Cần kiểm định giả thuyết: Trình độ học vấn không ảnh hưởng đến sự thay đổi lựa chọn công việc

*Kiểm định giả thuyết:*

*-* Sử dụng hàm chi2\_contigency của thư việc scipy.stats, ta thu được kết quả:

x^2: 165.65538908719 p\_val: 8.949047910389103e-35 df: 4

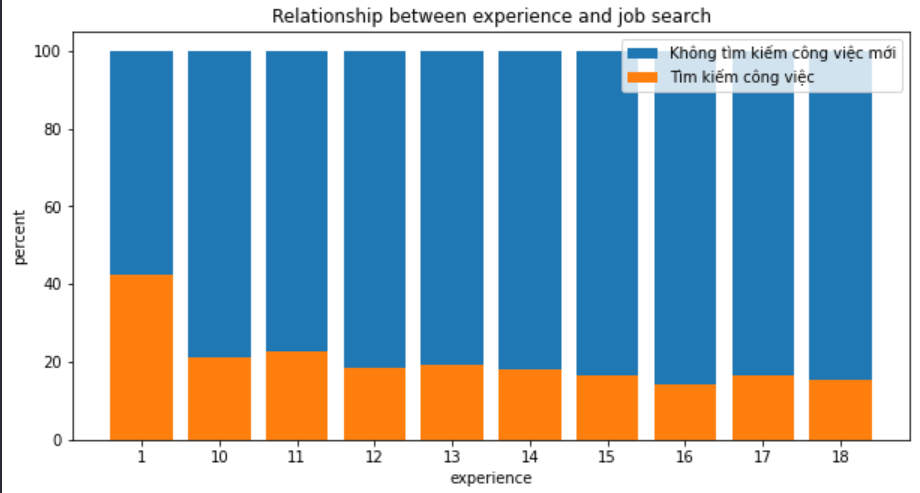
* Kết luận:

+ Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Có nghĩa là, việc biết trình độ học vấn của ứng viên cung cấp cho chúng tôi thông tin về việc liệu người đó sẽ tìm kiếm một công việc mới hay không.

+ Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

6.3. Số năm kinh nghiệm của ứng viên có ảnh hưởng đến việc thay đổi lựa chọn công việc không?

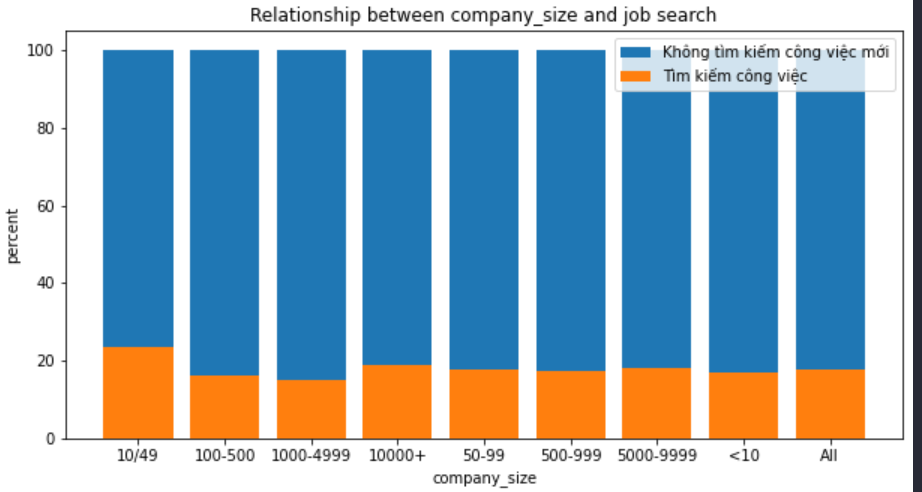
-Biểu đồ cột ghép thể hiện mối tương quan giữa biến experience và target



Kết luận: Kinh nghiệm làm việc càng ít, ứng viên càng có khả năng tìm kiếm công việc cao

**6.4. Quy mô công ty có ảnh hưởng đến việc lựa chọn thay đổi công việc không?**

**-** Sự tương quan giữa biến company\_size và target



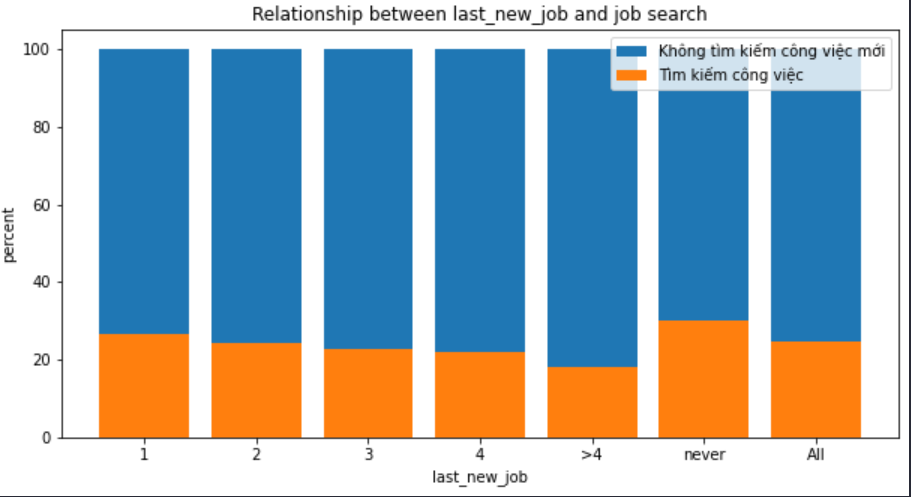
* Nhận xét:
  + Tỷ lệ ứng viên tìm kiếm một công việc mới khá nhất quán trên tất cả các quy mô công ty
  + Tỷ lệ người tìm kiếm một công việc mới gần như giống nhau ở tất cả các quy mô công ty.
  + Cần kiểm định giả thuyết: Quy mô công ty không ảnh hưởng đến việc thay đổi tỉ lệ tìm việc
* Kiểm định giả thuyết:
  + Sử dụng hàm chi2\_contigency, ta thu được kết quả:

x^2: 45.53175529315902 p\_val: 1.0781332047560822e-07 df: 7

* + Kết luận
    - Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Nghĩa là, việc biết quy mô công ty của ứng viên cung cấp cho chúng tôi thông tin về việc liệu người đó có tìm kiếm công việc mới hay không.
    - Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

**6.5. Sự khác biệt về số năm giữa công việc trước đây và công việc hiện tại có ảnh hưởng đến việc thay đổi công việc không?**

**-** Sự tương quan giữa last\_new\_job và target

****

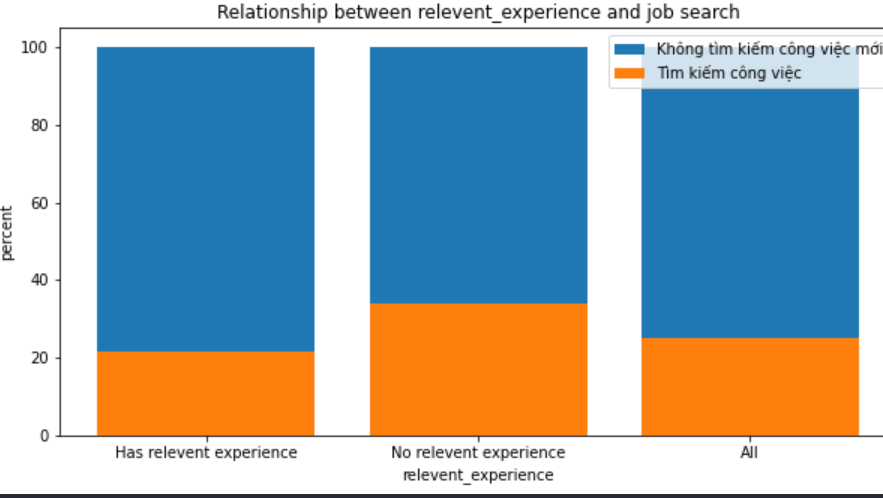
* Nhận xét:
  + Tỷ lệ ứng viên tìm kiếm một công việc mới khá nhất quán trên tất cả các danh mục last\_new\_job.
  + Tỷ lệ người tìm kiếm một công việc mới gần như giống nhau đối với tất cả các danh mục last\_new\_job.
  + Sự phân bố của những người tìm việc trong các danh mục khác nhau của last\_new\_job dường như giống nhau.
  + Nếu sự phân bố của những người tìm việc là như nhau trong tất cả các loại, điều đó cho thấy rằng hai biến này độc lập với nhau
  + Cần kiểm định giả thuyết: Danh mục last\_new\_job không ảnh hưởng đến việc thay đổi tỷ lệ người tìm việc, độ tin cậy 95%
* Kiểm định giả thuyết:
  + Sử dụng hàm chi2\_contigency, ta được:

x^2: 132.49948396906083 p\_val: 7.015169967001471e-27 df: 5

* + Kết luận:
    - Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Nghĩa là, việc biết last\_new\_job của ứng viên cung cấp cho chúng tôi thông tin về việc liệu người đó có tìm kiếm công việc mới hay không.
    - Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

**6.6. Kinh nghiệm tương đương của ứng viên có ảnh hưởng đến lựa chọn thay đổi công việc không?**

**-** Sự tương quan giữa relevent\_experience và target:



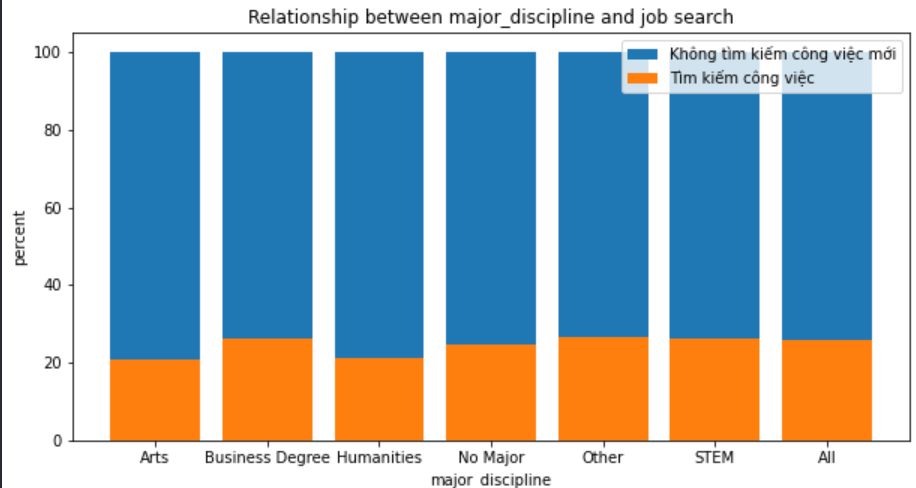
* Nhận xét:
  + Có vẻ như sự khác biệt về tỷ lệ người tìm việc ở hai loại là không đáng kể
  + Có vẻ như sự phân bố của những người tìm việc là ngang nhau trong hai loại và hai biến là độc lập
  + Chúng ta cần sử dụng thử nghiệm độc lập để chứng minh điều này.
* Kiểm định giả thuyết:
  + Sử dụng hàm chi2\_contigency, ta được:

x^2: 315.33857672021657 p\_val: 1.5006628411178982e-70 df: 1

* + Kết luận
    - Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Tức là, biết được mức\_thức\_thức\_của ứng viên sẽ cung cấp cho chúng tôi thông tin về việc liệu người đó có tìm kiếm một công việc mới hay không.
    - Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

**6.7. Chuyên ngành của ứng viên có ảnh hưởng đến việc thay đổi công việc hay không?**

**-** Sự tương quan giữa major\_discipline

- Nhận xét:

+ Có vẻ như sự khác biệt về tỷ lệ người tìm việc ở các hạng mục là không đáng kể.

+ Có vẻ như sự phân bố của những người tìm việc là ngang nhau trong các danh mục và hai biến là độc lập.

+ Chúng ta sử dụng thử nghiệm độc lập để chứng minh điều này.

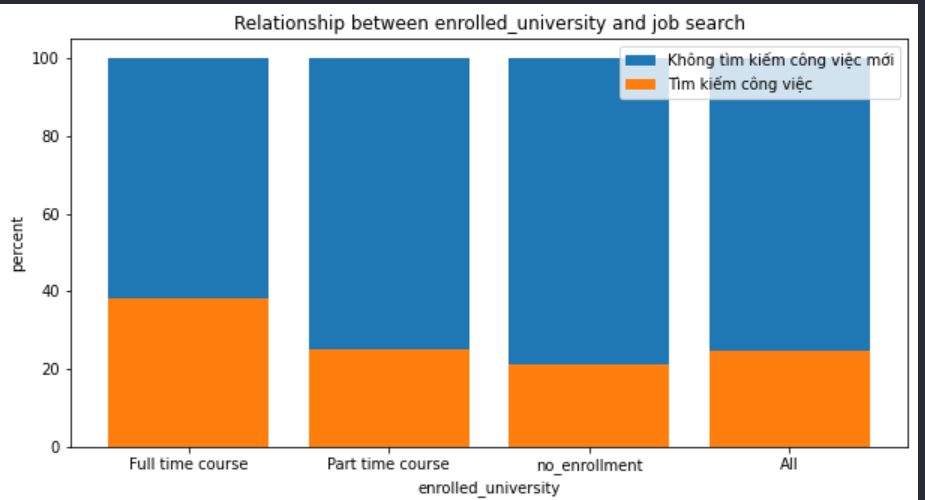
* Kiểm định giả thuyết:
  + Sử dụng hàm chi2\_contigency, ta được

x^2: 12.207121063439173 p\_val: 0.03205735800130323 df: 5

* + Kết luận
    - Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Nghĩa là, giả thuyết vô hiệu bị bác bỏ và hai biến phụ thuộc. Giá trị của p\_value gần với alpha. Điều này có nghĩa là chúng tôi có niềm tin thấp vào quyết định này
    - Ngoài ra, giá trị này gần với alpha, tức là có một mối quan hệ yếu giữa hai biến. Ngoài ra, giá trị alpha mong muốn là giá trị chung cho các thử nghiệm như vậy. Chúng ta đặt giá trị này và nó có thể thay đổi. Giá trị này có thể thay đổi theo mong đợi của chúng ta.

**6.8. Loại khóa học Đại học của ứng viên có ảnh hưởng đến việc lựa chọn thay đổi công việc không?**

**-** Sự tương quan giữa enrolled\_universty và target



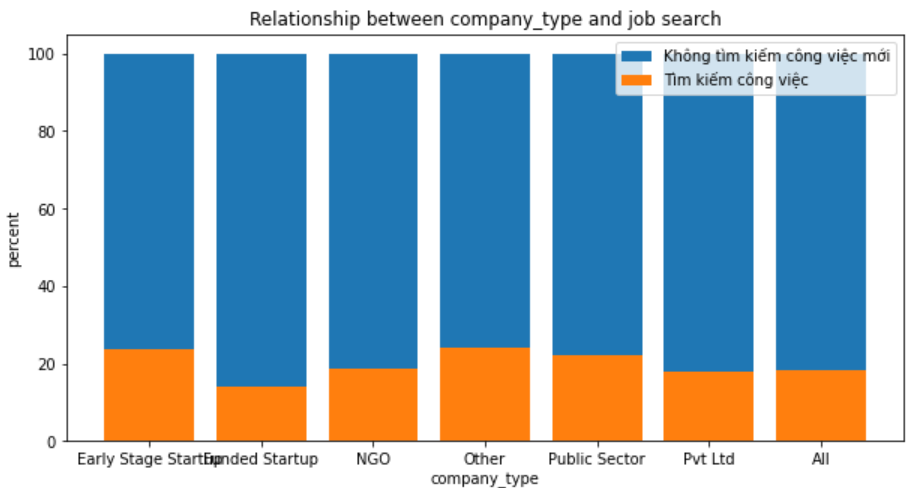
* Nhận xét: Ta cần kiểm tra tính độc lập của biến này
* Kiểm định giả thuyết
  + Sử dụng hàm chi2\_contigency, ta được:

x^2: 455.166762018575 p\_val: 1.4514211750080545e-99 df: 2

* + Kết luận
    - 1. Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Nghĩa là, việc biết được trường đăng ký của ứng viên sẽ cung cấp cho chúng tôi thông tin về việc liệu người đó có tìm kiếm công việc mới hay không.
    - Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

**6.9. Loại công ty có ảnh hưởng đến lựa chọn thay đổi công việc không?**

**-** Sự tương quan giữa company\_type



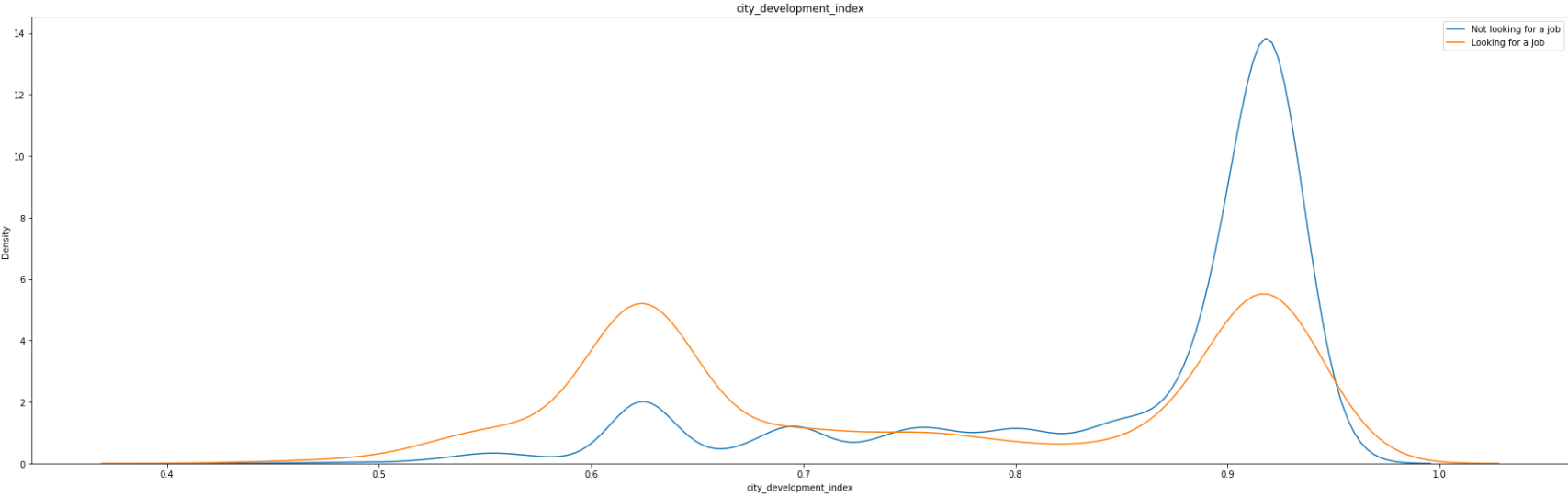
* Nhận xét: Ta cần kiểm tra tính độc lập của 2 biến
* Kiểm định giả thuyết:
  + Sử dụng hàm chi2\_contigency, ta được

x^2: 455.166762018575 p\_val: 1.4514211750080545e-99 df: 2

* + Kết luận
    - 1. Giá trị của p\_value nhỏ hơn alpha, tức là 0,05. Điều này có nghĩa là giả thuyết rỗng bị bác bỏ và hai biến phụ thuộc. Nghĩa là, việc biết loại\_công ty của ứng viên cung cấp cho chúng tôi thông tin về việc liệu người đó có tìm kiếm một công việc mới hay không.
    - 2. Giá trị p\_value quá xa so với alpha. Đó là, mối quan hệ giữa hai biến rất mạnh. Mối quan hệ giữa hai người gần như là một mối liên kết rất mạnh mẽ hơn là sự kết nối độc lập.

**6.10. Chỉ số phát triễn thành phố có ảnh hưởng đến việc lựa chọn thay đổi công việc không?**

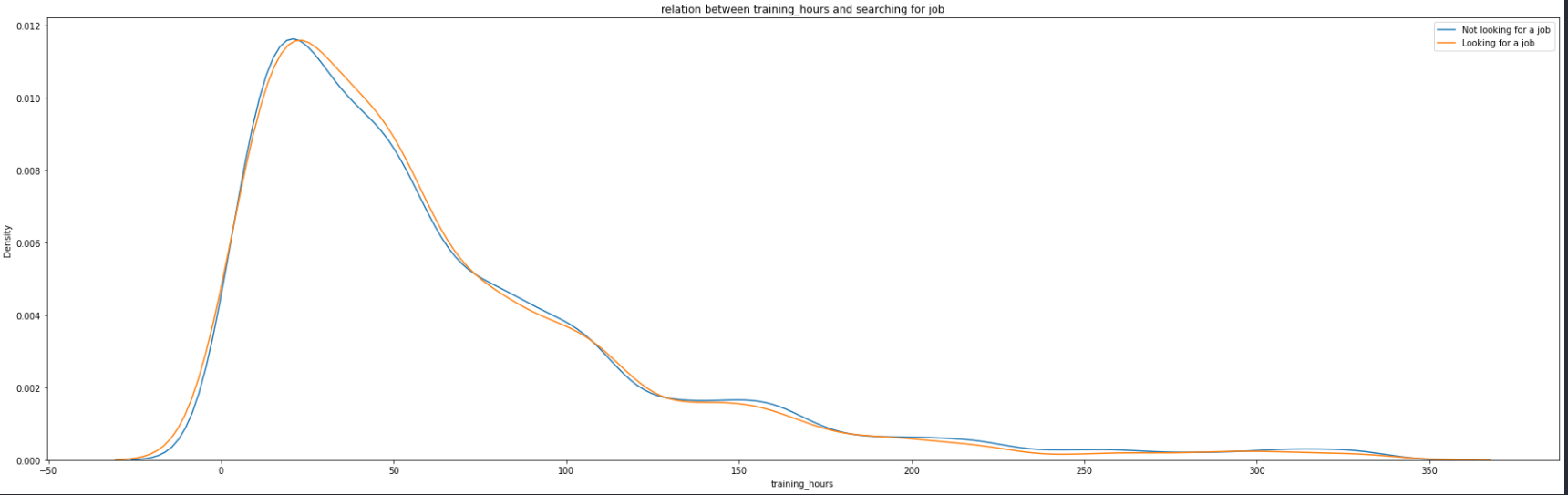
**-** Biểu đồ Kdeplot thể hiện sự tương quan giữa city\_development\_index và target



* Kết luận:
  + Hai bản phân phối không giống nhau. Có một phạm vi trong đó phân phối cho mục tiêu có giá trị bằng 0 cao hơn phân phối cho mục tiêu có giá trị bằng một. Có một phạm vi mà phân phối cho mục tiêu có giá trị là một cao hơn phân phối cho mục tiêu có giá trị bằng không.
  + Mục tiêu có giá trị là một có nghĩa là những người đang tìm kiếm một công việc mới. Sự khác biệt này giúp xác định giá trị mục tiêu của dữ liệu theo giá trị chỉ mục. Điều này có nghĩa là chỉ mục này có ảnh hưởng đáng kể đến nhãn dữ liệu và có mối quan hệ giữa nhãn và chỉ mục.

**6.11. Thời gian đào tạo có ảnh hưởng đến lựa chọn thay đổi công việc không?**

**-** Biểu đồ Kdeplot thể hiện sự tương quan giữa training\_hours và target

****

* Kết luận: Chúng ta có thể quan sát thấy rằng cả hai đường đều có xu hướng gần như giống nhau. Điều này có thể cho chúng tôi biết rằng giờ đào tạo không có tác động lớn đến việc tìm kiếm việc làm.