

1. GIỚI THIỆU BÀI TOÁN

1.1. Đặt vấn đề

Trong bối cảnh hiện đại, việc tuyển dụng nhân sự đóng vai trò quan trọng trong sự phát triển và thành công của các doanh nghiệp. Thị trường việc làm trong ngành CNTT luôn biến đổi nhanh chóng với sự thay đổi liên tục về nhu cầu kỹ năng, mức lương, và các yêu cầu tuyển dụng. Các ứng viên, nhà tuyển dụng và các nhà hoạch định chính sách đều cần hiểu rõ về xu hướng thị trường để đưa ra quyết định chính xác.

Bài toán "Phân tích và xử lý dữ liệu tuyển dụng ngành CNTT" tập trung vào việc sử dụng các kỹ thuật phân tích dữ liệu để khai thác thông tin từ các nguồn dữ liệu tuyển dụng. Từ đó đưa ra những phân tích, nhận định và dự đoán nhằm nâng cao chất lượng và hiệu quả của quá trình tuyển dụng đồng thời đưa ra các kỹ năng mà các ứng viên cần có để có thể ứng tuyển các công việc mà họ mong muốn.

Phân tích thị trường việc làm ngành CNTT cung cấp thông tin chi tiết và quan trọng, giúp trả lời nhiều câu hỏi liên quan đến tình hình và xu hướng việc làm.

Dưới đây là một số câu hỏi có thể giúp ta hiểu rõ hơn về thị trường việc làm trong ngành CNTT:

1. Nhu cầu lao động hiện tại là gì?

Những vị trí công việc nào đang có nhu cầu cao trong ngành CNTT? Các kỹ năng nào đang được yêu cầu nhiều nhất?

2. Mức lương và phúc lợi hiện tại ra sao?

Mức lương trung bình và các phúc lợi đi kèm cho từng vị trí, ngành nghề là bao nhiêu?

3. Địa điểm tuyển dụng phổ biến?

Những khu vực, thành phố hoặc vùng lãnh thổ nào đang có nhu cầu tuyển dụng cao nhất?

4. Yêu cầu kỹ năng và kinh nghiệm?

Các kỹ năng, trình độ học vấn và kinh nghiệm nào đang được các nhà tuyển dụng ưu tiên?

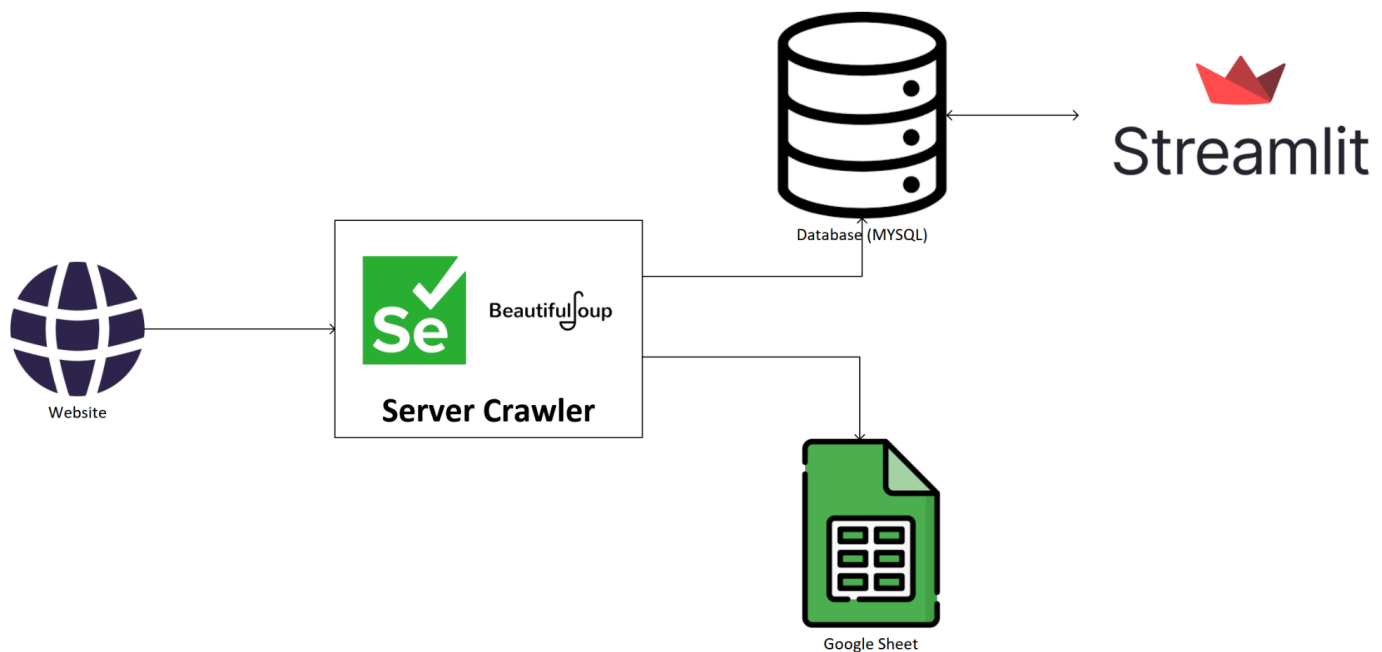
5. Thị trường lao động cho người mới tốt nghiệp?

Cơ hội việc làm và yêu cầu cho người mới tốt nghiệp, cũng như các ngành nghề phù hợp với họ.

6. Tình hình thất nghiệp và xu hướng?

Tỷ lệ thất nghiệp hiện tại, lý do chính dẫn đến thất nghiệp, và xu hướng thay đổi của tỷ lệ này.

1.2. Kiến trúc hệ thống



Sau khi xác định bài toán, với mục tiêu triển khai hệ thống đơn giản, dễ thao tác, giám sát, nhóm quyết định sẽ triển khai một máy tính chạy hệ điều hành Windows 10 với vai trò là máy chủ để thực hiện việc thu thập dữ liệu từ các trang web. Hệ thống sẽ sử dụng ngôn ngữ lập trình Python và các công nghệ như Selenium và BeautifulSoup để tự động duyệt web và thu thập thông tin.

Dữ liệu thu thập được sẽ được lưu trữ trong một cơ sở dữ liệu MySQL trên MySQL Server. Cơ sở dữ liệu này sẽ phục vụ cho quá trình xử lý và phân tích dữ liệu bằng Streamlit sau này.

Ngoài ra, dữ liệu cũng sẽ được sao lưu và hiển thị trên Google Sheet để dễ dàng theo dõi và chia sẻ thông tin với các thành viên trong nhóm. Điều này sẽ giúp tiện lợi trong việc theo dõi các chỉ số và thông tin cần thiết hàng ngày.

2. CÔNG NGHỆ SỬ DỤNG

2.1. Selenium

Selenium là một công cụ mạnh mẽ không chỉ trong kiểm thử tự động mà còn trong việc thu thập dữ liệu (web scraping) từ các trang web. Bằng cách mô phỏng các hành động của người dùng trên trình duyệt web, Selenium có thể truy cập, tương tác, và trích xuất dữ liệu từ các trang web động phức tạp mà các công cụ thu thập dữ liệu khác có thể gặp khó khăn.

- Ưu điểm của Selenium trong thu thập dữ liệu

- + Xử lý các trang web động: Tương tác với JavaScript: Selenium có khả năng xử lý các trang web có nội dung động được tải bởi JavaScript, giúp trích xuất dữ liệu mà các công cụ thu thập dữ liệu tĩnh không thể làm được.
- + Mô phỏng các tương tác phức tạp: Selenium có thể mô phỏng các hành động như cuộn trang, nhấp chuột, điền vào các biểu mẫu, và xử lý các cửa sổ bật lên (pop-up), điều mà các công cụ thu thập dữ liệu khác có thể không thực hiện được.
- + Đa nền tảng và đa trình duyệt: Hỗ trợ nhiều trình duyệt: Selenium hỗ trợ các trình duyệt phổ biến như Chrome, Firefox, Safari, Internet Explorer, Edge, và Opera. Điều này giúp bạn thu thập dữ liệu từ các trang web mà có thể yêu cầu một trình duyệt cụ thể để hiển thị đúng nội dung.
- + Đa nền tảng: Selenium có thể chạy trên các hệ điều hành khác nhau như Windows, macOS, và Linux, mang lại sự linh hoạt cho các dự án thu thập dữ liệu.
- + Tự động hóa các tác vụ lặp lại: Tự động hóa quy trình thu thập: Bạn có thể lập trình Selenium để tự động truy cập các trang web, điền vào các biểu mẫu, và trích xuất dữ liệu theo định kỳ, giúp tiết kiệm thời gian và công sức so với việc thực hiện thủ công.
- + Quản lý phiên làm việc: Selenium có thể lưu trữ và quản lý các phiên làm việc, giúp duy trì trạng thái đăng nhập hoặc xử lý các phiên làm việc phức tạp khi thu thập dữ liệu.
- + Dễ dàng tích hợp: Tích hợp với các ngôn ngữ lập trình: Selenium hỗ trợ nhiều ngôn ngữ lập trình như Python, Java, C#, và JavaScript, giúp dễ dàng tích hợp với các ứng dụng và quy trình hiện có của bạn.
- + Kết hợp với các công cụ khác: Bạn có thể kết hợp Selenium với các thư viện xử lý dữ liệu và phân tích như BeautifulSoup, Pandas, hoặc Scrapy để tạo ra các quy trình thu thập và xử lý dữ liệu hoàn chỉnh.
- + Xử lý dữ liệu phức tạp: Thu thập dữ liệu từ các phần tử động: Selenium có thể trích xuất dữ liệu từ các phần tử động, như các bảng dữ liệu được tải động hoặc các phần tử được ẩn/hiện dựa trên tương tác người dùng.

- + Xử lý các trang bảo vệ bằng CAPTCHA: Mặc dù không thể tự động giải quyết CAPTCHA, Selenium có thể dừng và chờ để người dùng giải quyết CAPTCHA trước khi tiếp tục, giúp thu thập dữ liệu từ các trang có bảo vệ CAPTCHA.
- + Cộng đồng lớn và tài liệu phong phú: Với cộng đồng người dùng và phát triển rộng lớn, bạn có thể dễ dàng tìm thấy tài liệu, hướng dẫn, và hỗ trợ từ các diễn đàn, blog, và trang web liên quan đến Selenium.
- + Cập nhật liên tục: Selenium liên tục được cập nhật và cải tiến để hỗ trợ các trình duyệt và công nghệ web mới nhất.

2.2. MySQL

MySQL là một hệ quản trị cơ sở dữ liệu quan hệ mã nguồn mở, phổ biến và được sử dụng rộng rãi trên toàn thế giới. Nó được phát triển bởi công ty MySQL AB, và hiện nay được sở hữu và duy trì bởi Oracle Corporation. MySQL là một phần của nhiều giải pháp công nghệ phổ biến, như LAMP stack (Linux, Apache, MySQL, PHP/Perl/Python).

- Ưu điểm của MySQL:

- + Mã nguồn mở: MySQL là phần mềm mã nguồn mở, nghĩa là bạn có thể sử dụng và chỉnh sửa nó miễn phí. Điều này giúp tiết kiệm chi phí, đặc biệt cho các doanh nghiệp nhỏ và các dự án cá nhân.
- + Hiệu suất cao: MySQL được thiết kế để xử lý một lượng lớn dữ liệu với hiệu suất cao. Nó hỗ trợ nhiều tính năng tối ưu hóa như caching, indexing, và phân vùng dữ liệu.
- + Dễ sử dụng: MySQL có cú pháp SQL dễ hiểu và dễ học. Ngoài ra, có nhiều công cụ quản lý cơ sở dữ liệu đồ họa (như phpMyAdmin, MySQL Workbench) giúp việc quản trị cơ sở dữ liệu trở nên dễ dàng hơn.
- + Bảo mật: MySQL cung cấp nhiều tính năng bảo mật như kiểm soát truy cập dựa trên vai trò, mã hóa dữ liệu, và hỗ trợ xác thực người dùng mạnh mẽ.
- + Khả năng mở rộng: MySQL có thể được sử dụng cho các ứng dụng nhỏ và lớn, từ các trang web cá nhân đến các ứng dụng doanh nghiệp lớn với hàng triệu giao dịch mỗi ngày.
- + Hỗ trợ tốt: Với cộng đồng người dùng lớn và sự hỗ trợ từ Oracle, MySQL có tài liệu phong phú, diễn đàn hỗ trợ và nhiều nguồn tài nguyên học tập trực tuyến.
- + Tích hợp dễ dàng: MySQL dễ dàng tích hợp với nhiều ngôn ngữ lập trình và công nghệ khác nhau như PHP, Java, Python, C#, Ruby, và nhiều hơn nữa.
- + Sao lưu và khôi phục dữ liệu: MySQL cung cấp các công cụ mạnh mẽ cho việc sao lưu và khôi phục dữ liệu, đảm bảo an toàn và liên tục cho dữ liệu của bạn.

- **Một số trường hợp sử dụng phổ biến của MySQL**
- + Web Applications: Các trang web và ứng dụng web như WordPress, Joomla, và Drupal sử dụng MySQL để lưu trữ dữ liệu.
- + Data Warehousing: MySQL có thể được sử dụng trong các hệ thống kho dữ liệu để phân tích và báo cáo dữ liệu.
- + Content Management: Nhiều hệ thống quản lý nội dung sử dụng MySQL để lưu trữ và quản lý nội dung người dùng

2.3. Streamlit

Streamlit là một thư viện Python mã nguồn mở cho phép bạn tạo các ứng dụng web tương tác một cách nhanh chóng và dễ dàng. Được thiết kế đặc biệt cho các nhà khoa học dữ liệu và kỹ sư học máy, Streamlit giúp biến các tập lệnh Python thành các ứng dụng web mà không cần phải có kiến thức về phát triển web.

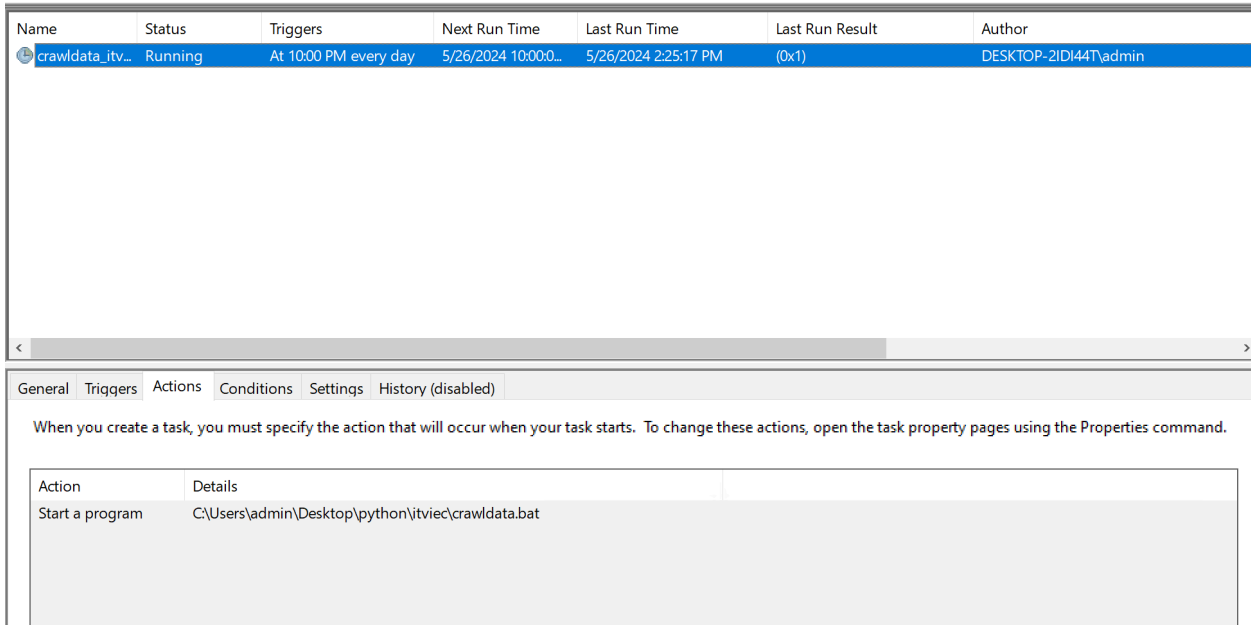
- **Ưu điểm của Streamlit:**
- + Dễ dàng sử dụng: Streamlit có cú pháp đơn giản và trực quan. Bạn chỉ cần vài dòng mã để tạo ra một ứng dụng web. Không cần phải lo lắng về HTML, CSS hay JavaScript.
- + Tương tác thời gian thực: Streamlit cho phép bạn tạo các ứng dụng web tương tác với khả năng cập nhật dữ liệu và giao diện người dùng trong thời gian thực. Bạn có thể thêm các widget như sliders, buttons, text inputs, và nhiều hơn nữa để người dùng tương tác trực tiếp với ứng dụng.
- + Tích hợp dễ dàng với các thư viện Python khác: Streamlit tích hợp tốt với các thư viện Python phổ biến như NumPy, Pandas, Matplotlib, Plotly, và Scikit-learn. Điều này giúp bạn dễ dàng hiển thị dữ liệu và các mô hình học máy trong ứng dụng của mình.
- + Triển khai nhanh chóng: Với Streamlit, bạn có thể triển khai ứng dụng của mình chỉ với một dòng lệnh. Streamlit cung cấp dịch vụ Streamlit Sharing cho phép bạn dễ dàng chia sẻ ứng dụng của mình trực tuyến.
- + Hiệu suất cao: Streamlit được thiết kế để tối ưu hóa hiệu suất. Các ứng dụng Streamlit chạy trên nền tảng Streamlit Server, cung cấp hiệu suất cao và khả năng mở rộng tốt.
- + Cộng đồng và tài liệu phong phú: Streamlit có một cộng đồng người dùng tích cực và nhiều tài liệu hỗ trợ, giúp bạn dễ dàng tìm kiếm giải pháp cho các vấn đề thường gặp và học hỏi từ các ví dụ và dự án của người khác.
- + Khả năng mở rộng: Streamlit cho phép bạn mở rộng các tính năng bằng cách sử dụng các thành phần tùy chỉnh và tích hợp với các API bên ngoài.

- **Một số trường hợp sử dụng phổ biến của Streamlit:**
- + Data Exploration and Visualization: Tạo các ứng dụng để khám phá và trực quan hóa dữ liệu một cách dễ dàng.
- + Machine Learning: Xây dựng các giao diện người dùng để hiển thị và tương tác với các mô hình học máy.
- + Dashboard: Tạo các bảng điều khiển tùy chỉnh để giám sát các chỉ số và KPI quan trọng.
- + Prototyping: Nhanh chóng tạo ra các nguyên mẫu cho các ứng dụng dữ liệu và kiểm tra các ý tưởng mới.

3. THU THẬP VÀ XỬ LÝ DỮ LIỆU

3.1. Thu thập dữ liệu

Để thực hiện việc thu thập dữ liệu theo lịch trình đã được thiết lập, nhóm sử dụng tính năng Task Scheduler của Windows 10 để tự động kích hoạt tiến trình vào 22h hàng ngày.



Tiến trình tự động này sẽ thực hiện bằng Python và gồm các bước sau:

- **Đăng nhập:** Trước tiên, tiến trình sẽ tự động đăng nhập vào trang web cần thu thập dữ liệu. Bước thực hiện này là cần thiết vì một số trang web yêu cầu đăng nhập để truy cập thông tin đầy đủ về một công việc.

```
# Tài khoản itviec
email = "rfd1894@gmail.com"
password = ""
service = Service(executable_path='C:/Users/admin/Desktop/python/itviec/chromedriver-win64/chromedriver.exe')

options = uc.ChromeOptions()
options.headless = False
driver = uc.Chrome(service=service, options=options)

# Login:
link_login = "https://itviec.com/sign_in"
driver.get(link_login)
email_input = driver.find_element(By.ID, "user_email")
email_input.send_keys(email)
password_input = driver.find_element(By.ID, "user_password")
password_input.send_keys(password)
sign_in_button = driver.find_element(By.XPATH, "//span[text()='Sign In with Email']")
sign_in_button.click()
```

- **Thu thập dữ liệu:** Sau khi đăng nhập thành công, tiến trình sẽ thực hiện thu thập toàn bộ dữ liệu của công việc. Tại bước này, dữ liệu sẽ được thu thập lưu thành các trường thông tin mà không cần xử lý, giữ nguyên dạng thô để tránh việc mất thông tin

quan trọng.

```
for i in range(40, 0, -1):
    link_web = "https://itviec.com/viec-lam-it?job_selected=none&page=" + str(i)
    driver.get(link_web)
    driver.implicitly_wait(10)
    print('Collecting Page' + f' {i}')
    page_source = driver.page_source
    soup = BeautifulSoup(page_source, 'html.parser')
    job_links = soup.find_all('a', {'data-search--job-selection-target': 'jobTitle'})
    href_list = []
    for link in job_links:
        href = link.get('href')
        href_list.append(href)
    for href in href_list:
        sleep(3)
        link_job = "https://itviec.com" + href
        driver.get(link_job)
        driver.implicitly_wait(10)
        page_source = driver.page_source
        soup = BeautifulSoup(page_source, 'html.parser')

        script_tag = soup.find('script', type='application/ld+json')

        if script_tag:
            script_content = script_tag.string
            job_info = json.loads(script_content)

            if "@type" in job_info and job_info["@type"] == "JobPosting":
                job_title = job_info["title"]
                job_salary = job_info.get("baseSalary", {})
                min_salary = job_salary.get("value", {}).get("minValue")
                max_salary = job_salary.get("value", {}).get("maxValue")
                employmentType = soup.find('span', class_='normal-text text-rich-grey ms-1').text.strip()
                company_link_a = soup.find('a', class_='hyperlink')
                company_link = company_link_a.get('href') if company_link_a else ""
                company_name = job_info["hiringOrganization"]["name"]
                job_posted_date = job_info["datePosted"]
                for location in job_info.get("jobLocation", []):
                    addressRegion = location.get("address", {}).get("addressRegion")
                    addressCountry = location.get("address", {}).get("addressCountry")
                job_skills = job_info["skills"]
                job_description = job_info["description"]
                job_start_index = job_description.find("The Job")
                your_skills_start_index = job_description.find("Your Skills and Experience")
                your_skills_end_index = job_description.find("Why You'll Love Working Here")
                job_description_job_part = job_description[job_start_index:your_skills_start_index]
                job_description_skills_part = job_description[your_skills_start_index:your_skills_end_index]
```

- **Lưu trữ dữ liệu vào Google Sheet và MySQL:** Trong quá trình thu thập dữ liệu, có thể xảy ra các vấn đề như lỗi code, lỗi máy chủ, hoặc bị giới hạn yêu cầu từ phía server của trang web. Do đó, cần có cơ chế để cập nhật dữ liệu mỗi khi có thông tin mới được thu thập. Dữ liệu sẽ được đẩy vào Google Sheet để dễ dàng quản lý và chia sẻ, cũng như lưu trữ trong cơ sở dữ liệu MySQL để phục vụ cho quá trình xử lý và phân tích sau này.


```

if job_posted_date == today:
    result = sheet.values().get(
        spreadsheetId=SPREADSHEET_ID,
        range=RANGE_NAME,
        majorDimension="ROWS"
    ).execute()

    values = result.get("values", [])
    next_row_index = len(values) + 1

    data_to_upload = [
        [job_title, href, min_salary, max_salary, employmentType,
        job_posted_date, company_name, company_link, addressRegion,
        addressCountry, job_skills, job_description_job_part,
        job_description_skills_part]
    ]

    request = sheet.values().update(
        spreadsheetId=SPREADSHEET_ID,
        range=RANGE_NAME + f"!A{next_row_index}",
        valueInputOption="RAW",
        body={"values": data_to_upload},
    )
    response = request.execute()

    print("Job data" + f" {job_title}"+" uploaded successfully!")

    cursor.execute("""
        INSERT INTO jobs (
            job_title, href, min_salary, max_salary, employmentType,
            job_posted_date, company_name, company_link, addressRegion,
            addressCountry, job_skills, job_description_job_part,
            job_description_skills_part
        ) VALUES (%s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s)
    """, (job_title, href, min_salary, max_salary, employmentType,
        job_posted_date, company_name, company_link, addressRegion,
        addressCountry, job_skills, job_description_job_part,
        job_description_skills_part))
    conn.commit()
    print("Job data" + f" {job_title}"+" inserted into MySQL successfully!")

```

Sau khi thu thập dữ liệu, nhóm em có được tập dữ liệu như sau:

job_title	salary_min	salary_max	employment_type	job_posted_date	company_name	city	country	skill
Software Architect DevOps, Project Manager			Tại văn phòng	4/8/2024	APEC Group	Hà Nội	VN	Software Architect, Project Manager, DevOps
Middle/Senior/Lead NodeJS Developer			Linh hoạt	4/8/2024	NFQ Asia	Hồ Chí Minh	VN	NodeJS, PostgreSQL, JavaScript
Junior UX Researcher Research Ops			Linh hoạt	4/8/2024	Ninja Van Tech Lab	Hồ Chí Minh	VN	UI-UX, English, Project Manager, Fresher Accepted
Middle-Senior Fullstack JS Developer NodeJS - React	2000	2500	Tại văn phòng	4/8/2024	Gigaron Technology	Hồ Chí Minh	VN	ReactJS, JavaScript, NodeJS
Senior UI/UX Designer	1000	2000	Tại văn phòng	4/8/2024	Fullerton Health	Hồ Chí Minh	VN	Designer, UI-UX, English
Senior QA Tester QA QC, Automation test	1000	2000	Tại văn phòng	4/8/2024	Fullerton Health	Hồ Chí Minh	VN	Tester, English, QA QC
Senior Java Dev - TNEX Finance - 3H121			Tại văn phòng	4/8/2024	Tập đoàn ROX (ROX)	Hà Nội	VN	Java, MySQL
DevOps Engineer Agile, AWS			Tại văn phòng	4/8/2024	Relationship Inc.	Hồ Chí Minh	VN	DevOps, Agile, AWS, Fresher Accepted
Project Manager			Tại văn phòng	4/8/2024	FuelCloud	Hồ Chí Minh	VN	Project Manager, Team Leader, English
Backend Developer PHP, C++, C#			Tại văn phòng	4/8/2024	Relationship Inc.	Hồ Chí Minh	VN	PHP, C#, C++
Java Spring Boot 2.5 Developer 1.500-2.000	1500	2000	Tại văn phòng	4/9/2024	BLUE DRAGON COM	Hồ Chí Minh	VN	Java, Spring
Software Engineer ReactJS, JavaScript, NodeJS	1000	3000	Tại văn phòng	4/9/2024	PAPAYA INSURTECH	Hồ Chí Minh	VN	ReactJS, JavaScript, NodeJS
Senior Automation Tester			Linh hoạt	4/9/2024	DEK Technologies	Hồ Chí Minh	VN	Automation Test, Java, Tester
Fullstack Developer .Net/ SQL/ VueJS			Tại văn phòng	4/9/2024	Gene Solutions	Hồ Chí Minh	VN	.NET, SQL, VueJS
Mobile Developer React Native	1000	2000	Tại văn phòng	4/9/2024	LOTTE Data Commu	Hồ Chí Minh	VN	React Native, JavaScript
02 Sr. QA QC Tester Manual / Automation	1000	1500	Tại văn phòng	4/9/2024	VUS	Hồ Chí Minh	VN	Tester, QA QC, Automation Test
QA Manager	2600	3000	Tại văn phòng	4/9/2024	Live Payments	Hồ Chí Minh	VN	QA QC, Agile, Manager
Solution Architect Java	1500	3000	Tại văn phòng	4/9/2024	ABBANK	Hà Nội	VN	Software Architect, Java, Spring
Senior Tester Manual/Automation/QA QC - TNEX - 3H120			Tại văn phòng	4/9/2024	Tập đoàn ROX (ROX)	Hà Nội	VN	QA QC, Agile, Tester
.NET Developer C#, ASP.NET 2 - 4 yrs			Tại văn phòng	4/9/2024	Pacific Technology	Hồ Chí Minh	VN	.NET, C#, ASP.NET

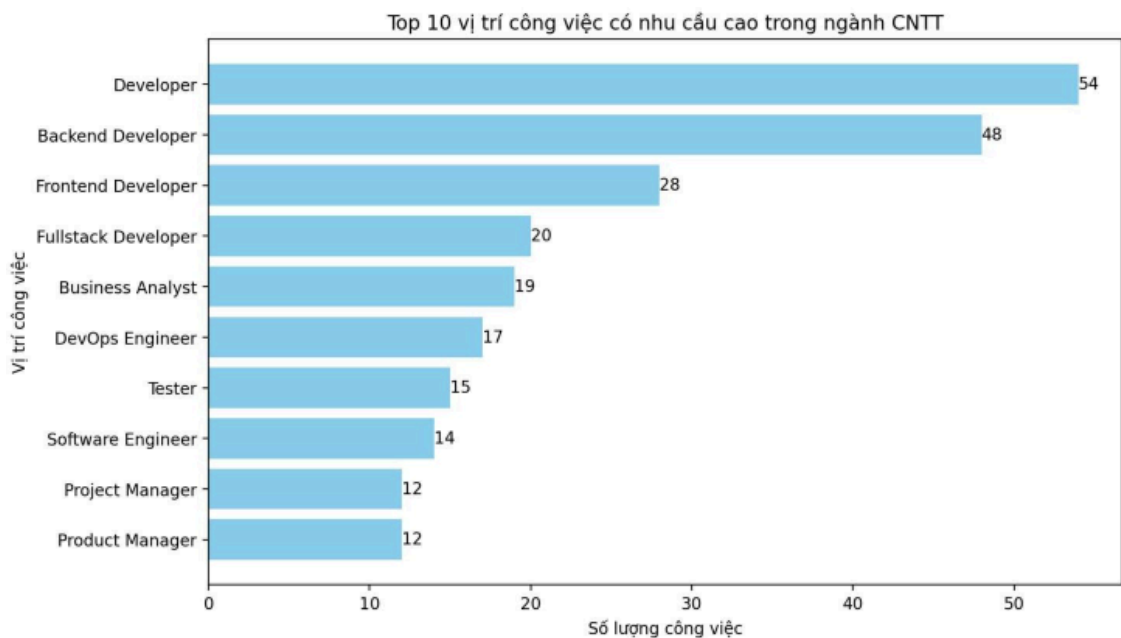
3.2. Xử lý dữ liệu

Để có thể trả lời những câu hỏi như đã đưa ra ở đầu bài, nhóm em cần làm sạch các bản ghi trong trường job_title và chuyển chúng thành tên vị trí công việc (position_name) và tên chức danh (job_title_separated). Nhóm em đã loại bỏ những thông tin không liên quan đến tên vị trí như thông tin về kỹ năng, lương, yêu cầu kinh nghiệm để có thông tin chính xác về tên vị trí. Chức danh của các vị trí sẽ bao gồm Fresher, Junior, Middle, Senior, Lead, Manager, Principal, Expert, All levels,...

job_title	position_name	job_title_separated
Software Architect DevOps, Project Manager	Software Architect	Manager
Middle/Senior/Lead NodeJS Developer	NodeJS Developer	Middle/Senior/Lead
Junior UX Researcher Research Ops	UX Researcher Research Ops	Junior
Middle-Senior Fullstack JS Developer NodeJS - ReactJS	Fullstack Developer	Middle- Senior
Senior UI/UX Designer	UI/UX Designer	Senior
Senior Java Dev - TNEX Finance - 3H121	Developer	Senior
.NET Developer C#, ASP.NET 2 - 4 yrs	.NET Developers	
Python Application Developer 1 - 3 years	Python Application Developer	
Senior Data Engineer Devops, AWS, SQL	Data Engineer	Senior
IT Server Administration Senior Officer AIX	IT Server Administration	Senior
DevOps Engineer Agile, AWS	DevOps Engineer	

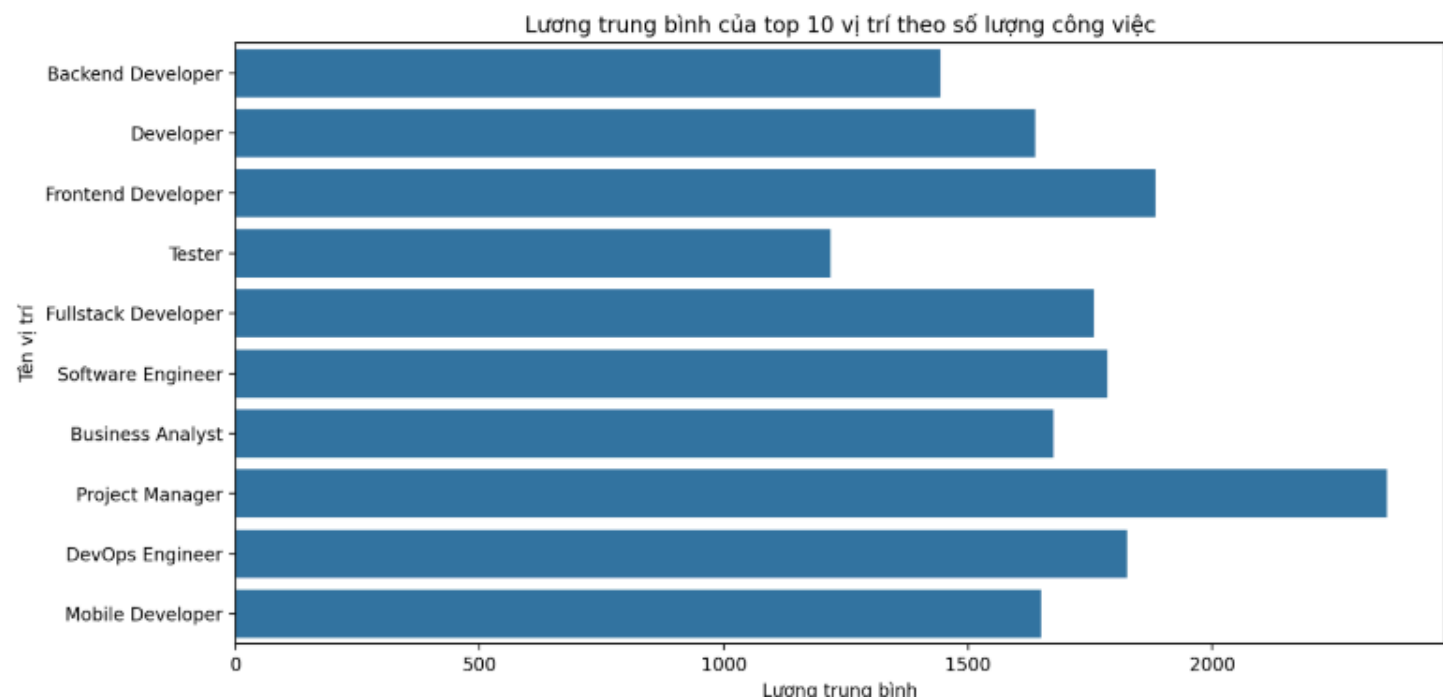
4. PHÂN TÍCH DỮ LIỆU

4.1. Nhu cầu lao động hiện nay



Những vị trí trong lĩnh vực phần mềm (như Developer, Business Analyst, Tester, Project Manager) vẫn đang rất hot và có nhu cầu cao ở thời điểm hiện tại thậm chí là trong tương lai xa hơn.

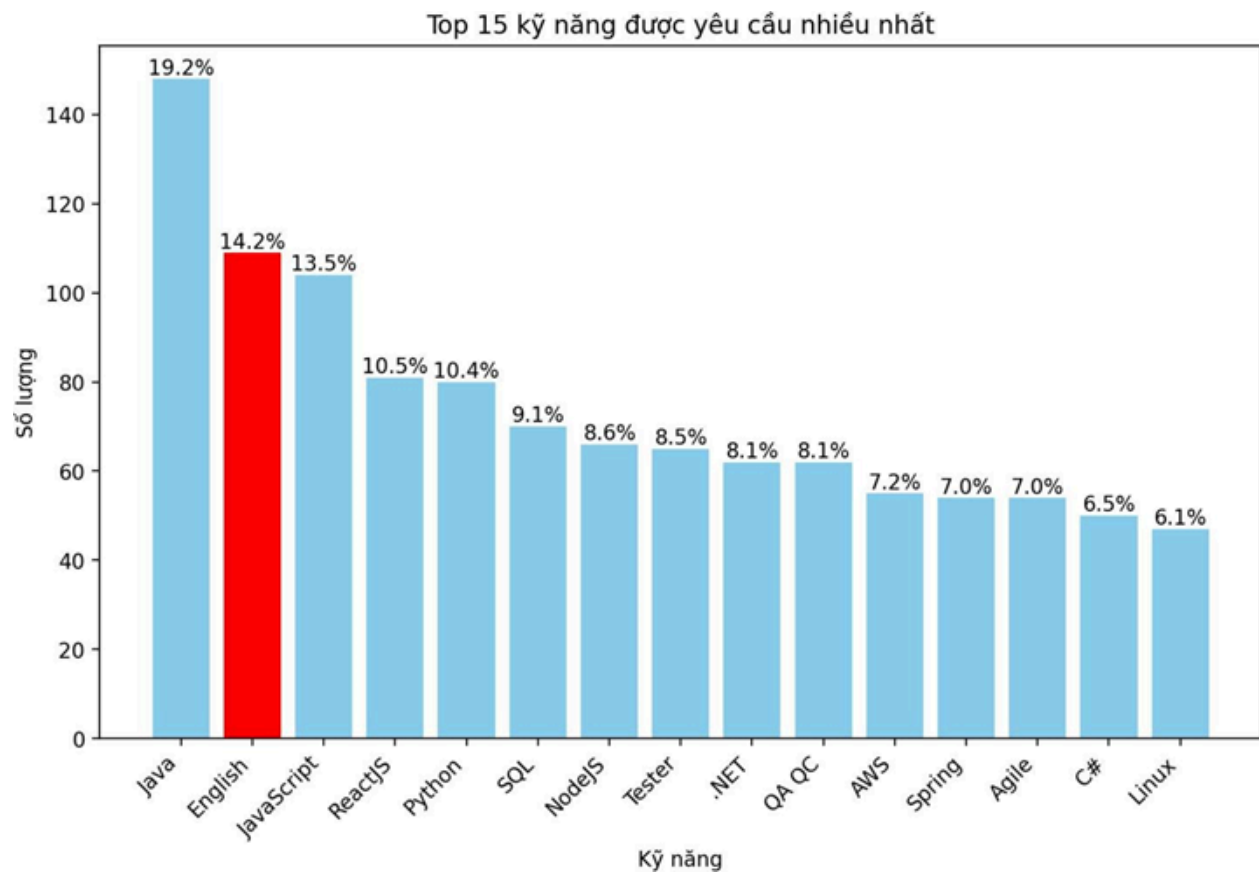
4.2. Mức lương và phúc lợi



position_name	job_count	avg_salary	cnt_senior	cnt_middle	cnt_junior
Backend Developer	17	1,443.0882	6	1	2
Developer	17	1,638.2353	6	3	3
Frontend Developer	9	1,883.3333	5	2	0
Tester	8	1,218.75	3	0	0
Fullstack Developer	7	1,757	2	1	0
Software Engineer	7	1,785.7143	2	0	1
Business Analyst	6	1,675	1	1	0
Project Manager	6	2,358.3333	1	0	1
DevOps Engineer	4	1,825	0	0	0
Mobile Developer	4	1,650	1	0	0

Nhìn chung, mức lương trung bình với các công việc trong lĩnh vực CNTT vẫn ở mức cao so mặt bằng chung các công việc khác cũng như GDP đầu người của Việt Nam.

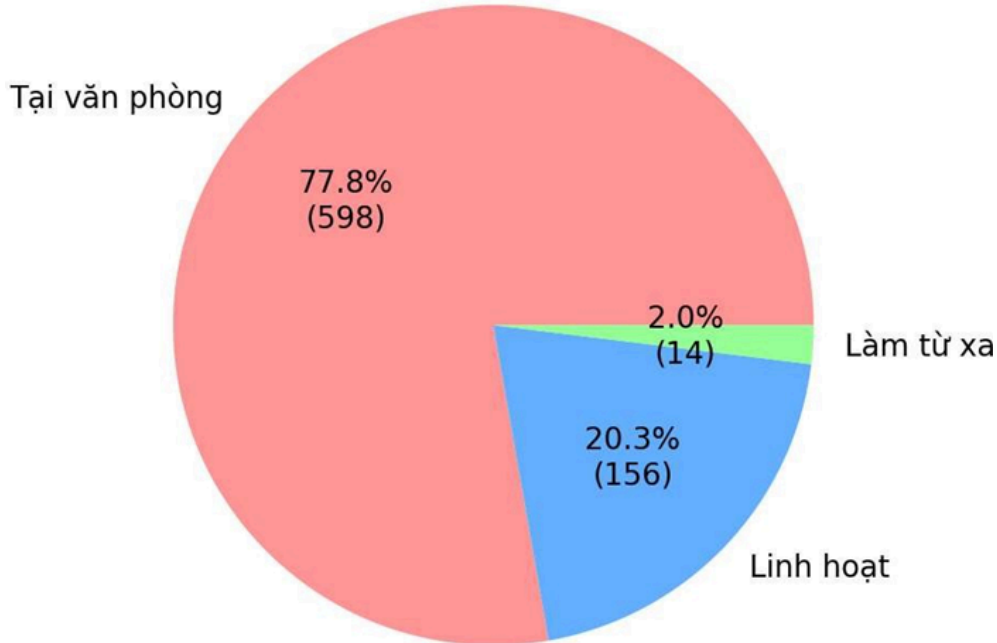
4.3. Những kỹ năng quan trọng, được các nhà tuyển dụng yêu cầu nhiều nhất



Những kỹ năng lập trình (như Java, JavaScript, Python,...) vẫn luôn được đặt lên hàng đầu. Cùng với đó, những trình độ Tiếng anh đang ngày càng trở nên quan trọng và trở thành điều kiện cần nếu ứng viên muốn nhận được công việc mà họ mong muốn.

4.4. Các loại hình công việc

Phân loại hình công việc

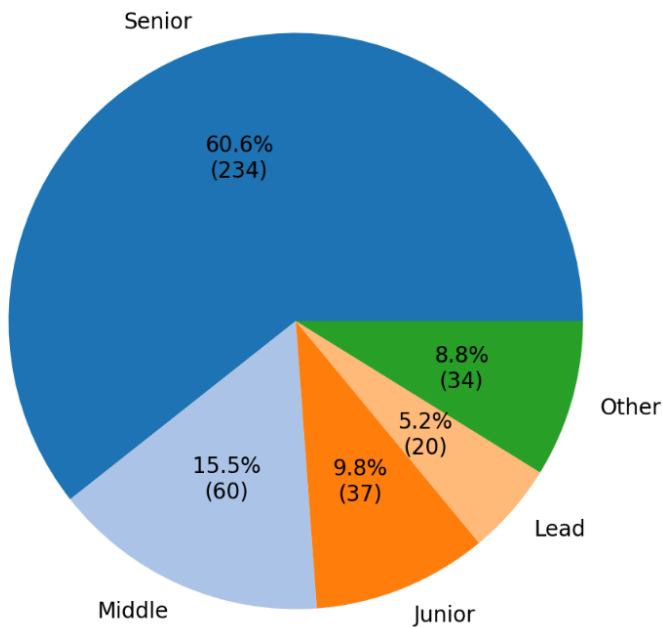
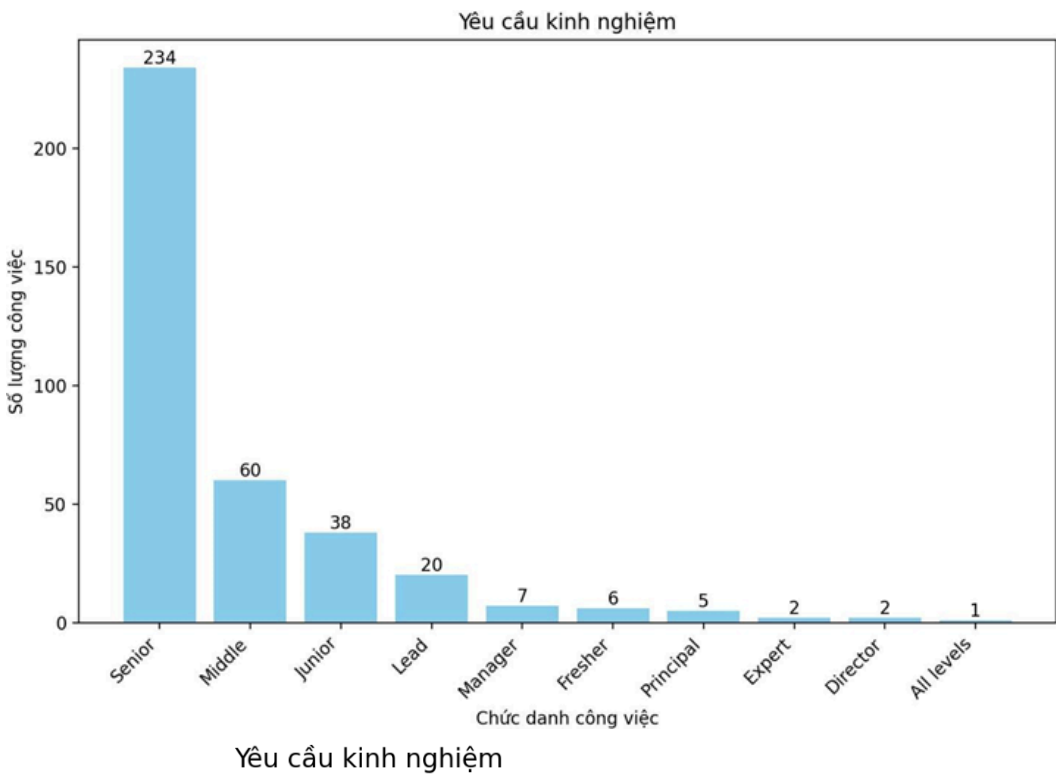


Từ biểu đồ tròn trên, có thể thấy rằng yêu cầu làm việc tại văn phòng chiếm tỷ lệ khá cao (77.8%).

Cùng với đó, có một lượng không nhỏ công ty cho phép nhân viên có thể linh hoạt làm việc tại văn phòng hoặc từ xa (chiếm 20.3%).

Hiện nay, mới chỉ có một số lượng rất ít công việc cho phép làm việc hoàn toàn từ xa (chỉ 2.0%). Điều này có thể đến từ việc các công ty cần sự giao tiếp trực tiếp giữa các nhân viên để tạo ra sự hiệu quả cao nhất trong công việc.

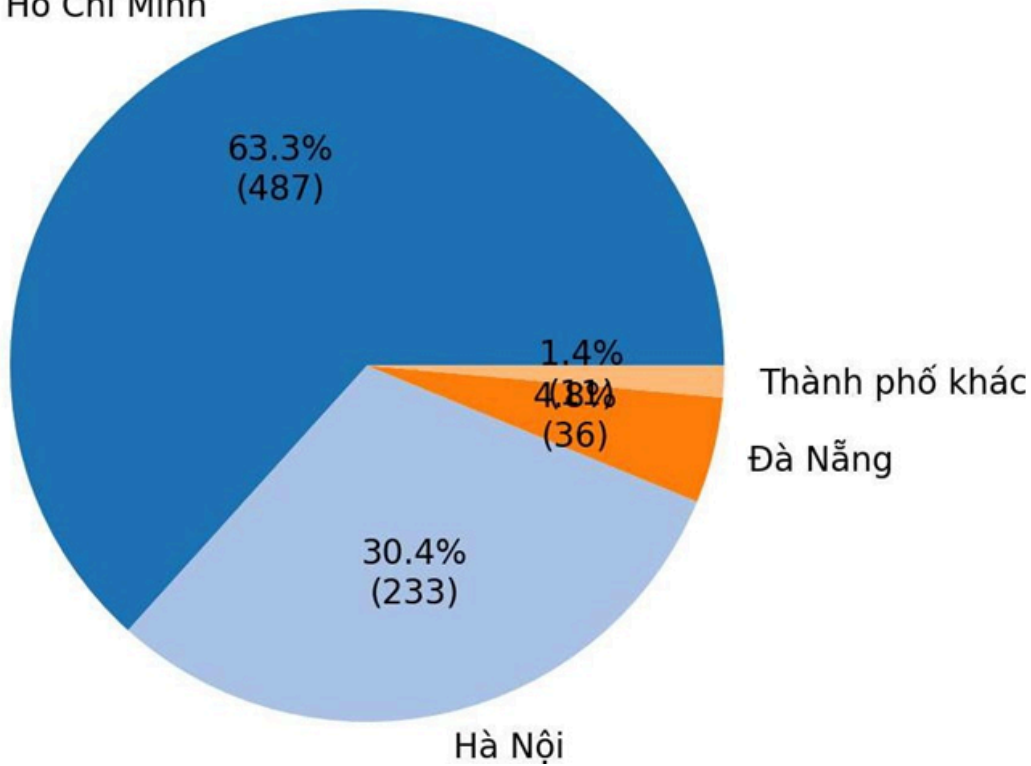
4.5. Yêu cầu về kỹ năng và kinh nghiệm



Hiện nay, các công ty thị trường lao động hiện tại đang thiên về việc tuyển dụng những ứng viên có nhiều kinh nghiệm, kỹ năng cao. Thị trường việc làm dành cho sinh viên mới ra trường hoặc những người ít kinh nghiệm là vô cùng cạnh tranh.

4.6. Địa điểm tuyển dụng phổ biến

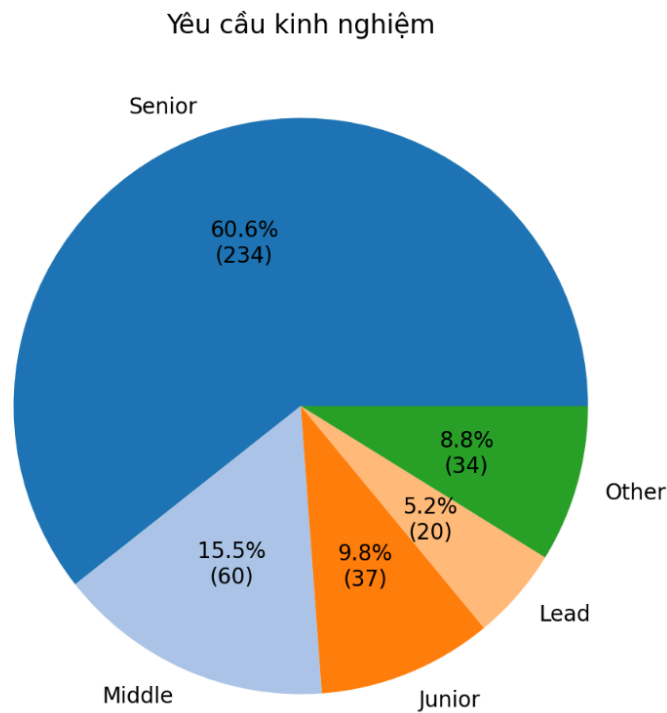
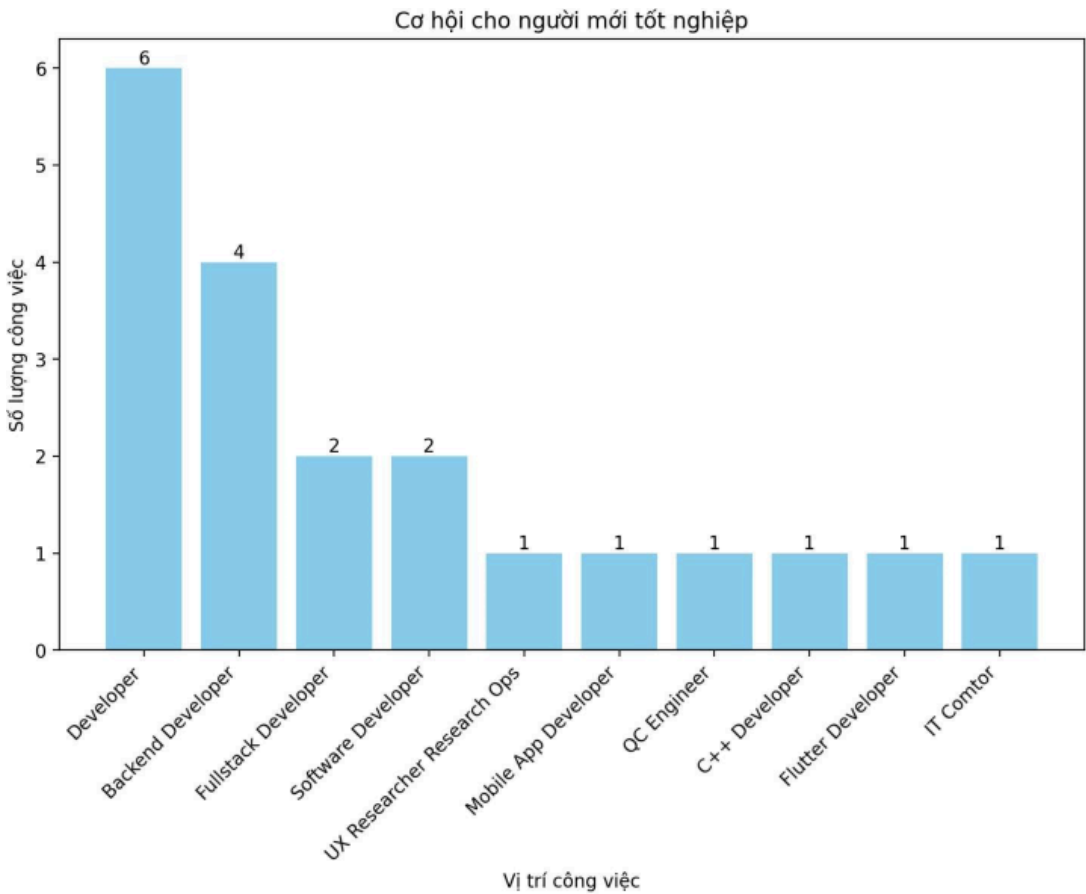
Hồ Chí Minh



TP. Hà Nội và TP. Hồ Chí Minh là nơi có nhu cầu tuyển dụng lao động ngành CNTT khi chiếm tới 93.7% tổng số công việc. Trong đó, TP. Hồ Chí Minh chiếm 63.3% và Hà Nội chiếm 30.4%.

Còn lại, TP. Đà Nẵng có 36 công việc chiếm tỷ lệ 4.9% còn lại là các tỉnh thành khác (1.4%).

4.7. Thị trường lao động cho người mới tốt nghiệp



Hiện nay, các công ty thị trường lao động hiện tại đang thiên về việc tuyển dụng những ứng viên có nhiều kinh nghiệm, kỹ năng cao. Thị trường việc làm dành cho sinh viên mới ra trường hoặc những người ít kinh nghiệm là vô cùng cạnh tranh.