AMERICAN UNIVERSITY OF BEIRUT

Career Dashboard for Students in Lebanon

by

Luna Baalbaki

Advisor(s)

Dr. Wissam Sammouri

Dr. Sirine Taleb

A Capstone Project

submitted in partial fulfillment of the requirements

for the degree of Master's in Business Analytics

to the Suliman S. Olayan School of Business

at the American University of Beirut

Beirut, Lebanon

August 2021

An Abstract of the Capstone Project of

Luna Baalbaki for Master's in Business Analytics (MSBA)

Title: Career Dashboard for Students in Lebanon

Selecting the right career is one of the most important decisions students have to make. With the increase in the number of career paths and opportunities, making this decision have become quite difficult. The purpose of this project is to do a job market analysis in Lebanon to provide information about different job trends in Lebanon up-to-date to Lebanese students, so they could form better career choices. Since many Lebanese students don't know which major to choose and which job position is highly demanded in Lebanon, this project guides them in their career choice.

The problem in Lebanon is that there's no available data about the job market in Lebanon. Thus,

I scrapped different websites to get information about job market trends in Lebanon and information about job professions as well. The design and methodology and methodology of this project is to scrape data about all job positions in Lebanon on LinkedIn, then clean and preprocess the data, turn it into a csv file then present it as visual graphs on a dashboard on Streamlit WebApp. Next, scrape data about job professions from O*net website to get information about different job professions. Then, clean and preprocess the data and present it as visual graphs on Streamlit WebApp. The major findings of this project were that the most demanded job position in Lebanon are manager and software engineer, the top hiring company is Toptal and Agoda, the city with the most job openings is Beirut. This project doesn't only benefit Lebanese students and help them choose the right career for them, but also, it helps Lebanon's economy. When worker's skills and qualifications fit demanded jobs in the country then that increases productivity and boosts economic growth.

Table of Contents

1. Introduction

In today's world choosing the right career is the toughest decision. Today many students are confused about their future. They do possess some skills but they are not able to identify their abilities and a proper domain. Different people suggest different career options but at last, the student has to select their career. Whether you're a student looking for career and college options or an adult looking for a new career path, this career guide capstone project is an easy tool for Lebanese students to explore job market trends in Lebanon and guide their career choice.

One of the problems that Lebanese students are facing is that they don't know what are the jobs that are most demanded in their country at the current year. Many Lebanese students are graduating with skills that don't match jobs available in Lebanon, and they end up either working in a profession that isn't what they wanted, immigrate or unemployed. This leads to an increase in unemployment rate, and loss in talent. All of these negatively affect the economy.

The objective of this project is to help students in Lebanon in choosing their profession through creating a WebApp that allows students to explore job market trends in Lebanon and guide their career choice. Also, to provide them with information about different professions, so that they could be more aware of tasks, skills and detailed work activities for the professions they are considering.

There are many available reports that are about labor market in Lebanon, but none of them is recent. In the "Labor Market in Lebanon" report that was issued by Statistics in Focus, employment by sector was shown. Even though this might be useful for students to know what are the sectors that are highly demanded, but it is very broad and the study has been done in 2011. After Lebanon's revolution,

financial crisis, political instability and pandemic employment by sector has definitely changed. Thus, these kinds of reports don't show the right job market trends in Lebanon that's in 2021.

At first, I thought about different approaches to provide recent job market analysis to Lebanese students. One of the approaches was to do a survey that would reach many Lebanese who are working in Lebanon and collect information about their jobs, industry they are working in and so on, then analyze the data that I collected. However, I realized that this information wouldn't be about the new job openings in Lebanon, and it's very hard for my survey to reach a lot people and to be a representative sample of the Lebanese population.

Another approach was to collect job market information from different reports that were done in Lebanon. However, this approach wouldn't be very effective because the data and trends that I will collect will be from past years and not from this year, because no recent study has been made on job market analysis in Lebanon.

Therefore, I decided to scrape data from LinkedIn and O*net website and I created a WebApp that portrays the job market trends in Lebanon in 2021, and provides information about different professions. This WebApp could be updated monthly by me. Thus, students would be always aware of the job market trends that's changing in Lebanon.

2. Background and Related Work

Many studies have shown some trends in job market in Lebanon. According to the German Cooperation, Lebanon's main economic sector in 2017 is the service sector (72.5 per cent of GDP), followed by industry (22.6 percent) and agriculture (4.9 per cent). Real estate is the largest service industry, contributing for an average of 13.7 percent of GDP from 2004 to 2016. Wholesale and retail (13.4 percent of GDP), public administration (9.4%), and financial services (7.3 percent) follow. These statistics could inform students who are in Lebanon on what are the most productive sectors that contribute well to the Lebanese economy (Knobloch, 2018)[1].

According to the Council of Scientific and Industrial Research (CSIR), over 40% of students are confused about their career options. Many of these students are confused because they are not informed about the exact work description of professions. This may lead to poor career choices and subsequent employment in fields that are not suited to them, lowering human resource productivity (Chaudhary, 2019)[2].

There is a major skills mismatch in terms of vacancies and the supply of highly qualified graduates. For finding out about open opportunities and labor market analyses, there should be an organized, centralized, and transparent platform that is continually updated (Chaudhary, 2019)[2]. Between December 2014 and February 2015, a total of 2,791 job openings covering Beirut, Mount Lebanon and North Lebanon were collected from local newspapers and websites (ILO, 2017)[3].

[1] Claudia Knobloch, (2018), German Cooperation, Employment and Labor Market Analysis in Lebanon.
[2] Chaudhary, D., Prajapati, H., Rathod, R., Patel, P., & Gurjwar, R. K. (2019).
Student Future Prediction Using Machine Learning.

[3] ILO, (2017), Labour Market Information Review and Analysis

These research studies could inform Lebanese students on what are the sectors are that are economically most productive, some trends in the working force… but it doesn't help students be aware of the sectors that have most job openings. Also, these studies are all done from previous years, no study that has been done and published this August 2021. They collect data over years and analyze them and then publish them, which by the times they publish their research paper, the job market trends might change. In addition, there's no study about the job market trends in Lebanon that could be updated monthly.

My approach that I took in my capstone project is that I analyzed more detailed information about job trends in Lebanon that could better help Lebanese students. Such as, job positions, companies and locations. In addition, in my project students could filter the graphs based on their wants, which makes it easier to students to get information, rather than reading lots of research papers about job trends in Lebanon.

Another way on how my project is different than the other studies, is that my project's data will be updated monthly by me. In that sense, my results will always be recent and accustomed to the crises that's happening on a monthly basis. Job market trends from 2017 in Lebanon are no more accurate because of the multiple crises that Lebanon is facing. Therefore, students can't rely on not updated job market trends because they are changing constantly.

3. Methodology

There are two main sources that I got my data from. The first one is LinkedIn and the second one is O*net. From both websites I had to scrape the data, clean it and transform it into csv files.

3. A) Scrapping Job Details in Lebanon from LinkedIn

Scrapping data from LinkedIn isn't very straight forward since LinkedIn puts lots of bots that not allow you to scrape data.

In this section, I will be going over how extracted data from LinkedIn about jobs positions (job title, company, location and date) using Selenium and other libraries. I've added the notebook, Streamlit link in my GitHub account @Luna Baalbaki.

The first step is to setup a WebDriver for Selenium on your computer. This allows Selenium to manipulate a browser. I used the Chrome WebDriver by updating Google Chrome and downloading the corresponding version from Chrome's WebDriver site.

Next, I used my WebDriver. We can call a GET request to LinkedIn by passing a string containing our search query. I used this link
https://www.linkedin.com/jobs/search?location=Lebanon&geoId=101834488&trk=public_jobs_jobs-search-bar_search-submit&currentJobId=2682763972&position=24&pageNum=6
which goes to LinkedIn jobs and searches for all the job positions available in Lebanon. From this page alone, there are all sorts of things we can scrape.

The main library that I will use is selenium, to be able to get information from the link that I want to scrape. Then I send a request to the browser, which acts as a hub for the requests to be passed on to the browser. The browser driver will pass the request to the browser and complete the actions.

My target from this notebook is to scrape data about job positions in Lebanon.

The URL that I'm scarping the data from is:

https://www.linkedin.com/jobs/search?keywords=&location=Lebanon&geoId=101834488&trk=public_jobs_jobs-search-bar_search-submit&position=1&pageNum=0

After running the browser, you are going to be directed to a LinkedIn page (shown below as image 1). On that page, scroll down and keep on clicking on "more jobs" to reach the last job position. In order to scrape all the jobs that are on the page. Then run the python codes (all codes are found in notebook which is found on my GitHub account @lunabaalbaki).

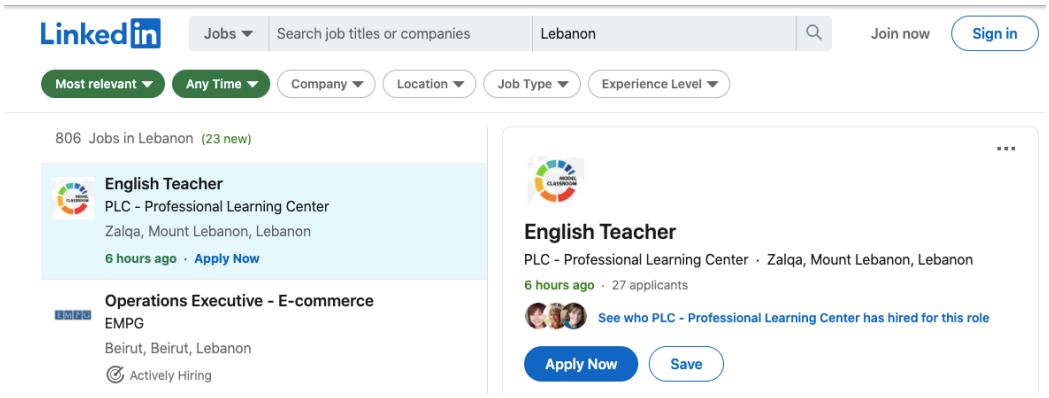


Image 1

LinkedIn page about jobs in Lebanon

This page will be the page that we will get our data from.

Next, we need to know how many jobs we have found through this search. Therefore, in order to scrape out of this link we need to find the HTML name of what we want to scrape.

Next, for scraping the name of job positions (title), right-click and click on inspect. Then, hover the mouse copy the name of the class of the “job title”. Then, paste the name in the python notebook code to get the information for the job title. To get the job title for all the jobs in the page, do a list and a for loop to get the job title for all the jobs in the page.

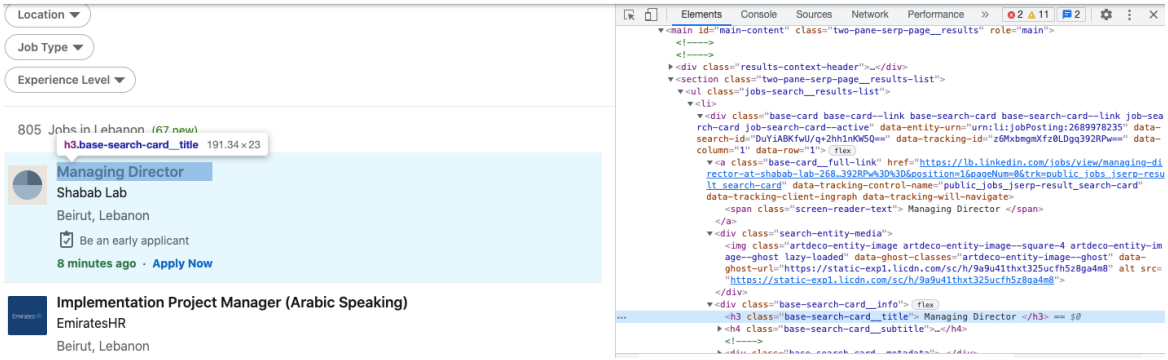


Image 2

LinkedIn page about scrapping jobs in Lebanon

Then, for scraping the name of the company, right-click and click on inspect. Then, copy the name of the class of the “company”. Then, paste the name in the python code to get the information for the company name. To get the company name for all the jobs in the page, do a list and a for loop.

Then, for scraping the location, right-click and click on inspect. Then, copy the name of the class of the “location”. Then, paste the name in the python code to get the information for the location. To get the location for all the jobs in the page, do a list and a for loop.

Then, for scraping the date, right-click and click on inspect. Then, copy the name of the class of the "date". Then, paste the name in the python code to get the information for the date. To get the date for all the jobs in the page, do a list and a for loop.

Right-click on your mouse, select 'inspect'. Click the 'Elements' and you will find the HTML codes. Then, copy the class name that is found in the HTML section and paste in the Jupyter notebook and run the code, you'll get the number of jobs you are going to scrape.

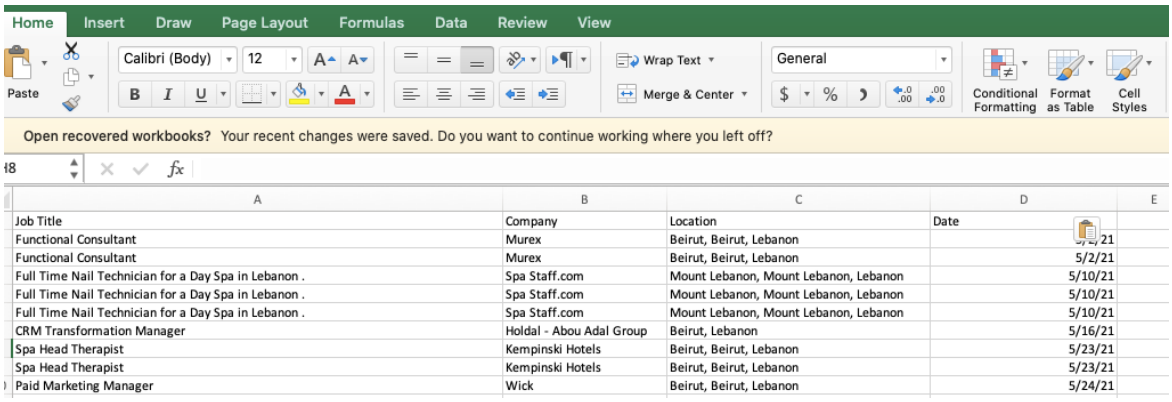
To combine job title, company name, location and date we are going to create a DataFrame for further analysis. The output of the DataFrame is shown below in image 4.

	Job Title	Company	Location	Date
0	HR and Purchasing Officer	Ecoline Services s.a.l	El Mkalless, Mount Lebanon, Lebanon	2021-08-20
1	Operations Management Intern	numoo	Beirut, Lebanon	2021-08-15
2	Operations Executive - E-commerce	EMPG	Beirut, Beirut, Lebanon	2021-08-19
3	Business Partner	Insurance Business America	Beyrouth, Beirut, Lebanon	2021-08-20
4	Chief Operating Officer	Yaduna WHHC	Baabda, Mount Lebanon, Lebanon	2021-08-02
5	Director General	St.John technical institute	Zgharta, North, Lebanon	2021-08-03
6	Business Development Lead	Siren Associates	Beirut, Lebanon	2021-08-20
7	Supply and Logistics Assistant	UNV LTD	Beirut, Beirut, Lebanon	2021-08-16

Image 4

DataFrame of job information scrapped from LinkedIn

Then, we are going to transform this DataFrame into a CSV file, in order to be able to clean it and analyze it using python. A snippet of the CSV file is shown in image 5.



Home Insert Draw Page Layout Formulas Data Review View

Paste | Calibri (Body) | 12 | Wrap Text | General | Conditional Formatting | Format as Table | Cell Styles

Open recovered workbooks? Your recent changes were saved. Do you want to continue working where you left off?

	A	B	C	D	E
	Job Title	Company	Location	Date	
	Functional Consultant	Murex	Beirut, Beirut, Lebanon	5/2/21	
	Functional Consultant	Murex	Beirut, Beirut, Lebanon	5/2/21	
	Full Time Nail Technician for a Day Spa in Lebanon .	Spa Staff.com	Mount Lebanon, Mount Lebanon, Lebanon	5/10/21	
	Full Time Nail Technician for a Day Spa in Lebanon .	Spa Staff.com	Mount Lebanon, Mount Lebanon, Lebanon	5/10/21	
	Full Time Nail Technician for a Day Spa in Lebanon .	Spa Staff.com	Mount Lebanon, Mount Lebanon, Lebanon	5/10/21	
	CRM Transformation Manager	Holdal - Abou Adal Group	Beirut, Lebanon	5/16/21	
	Spa Head Therapist	Kempinski Hotels	Beirut, Beirut, Lebanon	5/23/21	
	Spa Head Therapist	Kempinski Hotels	Beirut, Beirut, Lebanon	5/23/21	
	Paid Marketing Manager	Wick	Beirut, Beirut, Lebanon	5/24/21	

Image 5

CSV file of job information scrapped from LinkedIn

3. B) Cleaning and Preprocessing the Data that was Scrapped from LinkedIn

After transforming the data into a csv file, there was a lot of duplications, unknown characters, Arabic words, unnecessary digits, punctuations… Before visualizing the data, I had to clean and preprocess the data.

The first thing that I did is that I dropped all duplicates in that are the same across all columns. I made sure that they were dropped across all columns not at each single column alone because, in the "location" column for example, of course they are going to be duplications. Therefore, I only dropped the rows that are the same across all 4 columns.

Next, in the "location" column, I had to group locations that are the same but written differently together. For example, some jobs were written that they are located in "Beirut, Lebanon" and other jobs were written that they are located in "Beyrouthe, Lebanon". Since these two locations are the same, I had to group them. Therefore, using NLP and regex I found all locations that are in Beirut

and changed their location name to “Beirut, Lebanon”. I did the same for Mount Lebanon, North Lebanon, Beqaa, South Lebanon…

I transformed the “Date” column to months only instead of day-month-year for better visualization. Next, I focused on the “Company Name” column and I started cleaning it using regex. I removed URLs, unnecessary characters, punctuations, Arabic words and digits.

1	A2P Account Coordinator	Monty Mobile
2	Âæøê≥Ä„Éû„Éç„É º„ÇΠ„É£„É º	Movenpick
3	Academic Computer Center Su	Lebanese Americ

Image 6

Pre-cleaning the data

After cleaning, the data is now ready for visualization and extracting information from.

Scrapping Information about Professions

Data about professions were scrapped from O*net website. O*net website provides detailed information about each profession. I used same scrapping methods as shown above for scrapping information from the O*net and transform it into CSV files.

The O*net website that I want to scrape is shown below in image 7.

Browse by O*NET Data

Skills — Active Learning Save Table (XLS/CSV)

Understanding the implications of new information for both current and future problem-solving and decision-making.

Sort by: Importance ▲	Level	Code	Occupation
78	75	19-1021.00	Biochemists and Biophysicists
78	70	29-1229.05	Preventive Medicine Physicians
78	66	19-1041.00	Epidemiologists Bright Outlook
78	59	29-1171.00	Nurse Practitioners
75	79	19-2012.00	Physicists
75	75	19-1029.02	Molecular and Cellular Biologists
75	68	19-2011.00	Astronomers
75	68	11-1011.00	Chief Executives

Image 7

*Skills of Professions from O*net website*

I got the CSV file for every "category skill" and "skill" and I concatenated them all together along with their importance into one CSV file as shown below in image 7.

	A	B	C	D	E	F	G
	Skill	**Category of Skill**	**Importance**	**Level**	**Code**	**Profession**	**Importance1**
	Active Listening	Basic Skills	74	57	13-2011.00	accountant	74
	Reading Comprehension	Basic Skills	74	58	13-2011.00	accountant	74
	Critical Thinking	Basic Skills	72	55	13-2011.00	accountant	72
	Speaking	Basic Skills	70	55	13-2011.00	accountant	70
	Writing	Basic Skills	69	52	13-2011.00	accountant	69
	Judgement and Decision Making	Systems Skills	63	54	13-2011.00	accountant	63
	Complex Problem Solving	Complex Problem Solving	61	51	13-2011.00	accountant	61
	Active learning	Basic Skills	58	50	13-2011.00	accountant	58
	Time Management	Resource Management Skills	56	48	13-2011.00	accountant	56
	Systems Analysis	Systems Skills	55	52	13-2011.00	accountant	55
	Persuasion	Social Skills	53	47	13-2011.00	accountant	53
	Coordination	Social Skills	52	46	13-2011.00	accountant	52

Image 8

*CSV file of Skills scrapped from O*net Website.*

I also, wanted to scrape tasks and detailed work activities for every profession. So, I did the same steps that I did for the skills.

3. C) Streamlit

Streamlit is a WebApp where you can visualize your data. In this project I used Streamlit to graph my data and extract insights from it. This WebApp makes it easy for my audience to check the visuals that I will create from my data, insights, interact with my graphs and filter based on their interests.

In my Streamlit WebApp I did three pages. The first page is the “Home” page where it has an overview of my Streamlit Webapp and a navigation button. The next page is about “Job Market Analysis in Lebanon”, it contains visuals from the data that I scrapped from LinkedIn. The third page “Information about Professions” contains detailed information about different professions.

4. Results

After scrapping the data and cleaning it, I have used PowerBi at first to visualize my data and to know what graphs to I need to code on Streamlit app. On PowerBi I was able to graph the following:

- Jobs Based on Location
- Top Hiring Companies
- Top Demanded Jobs
- Months the Data Contains

After visualizing these graphs, I tried different types of graphs. Such as pie graphs, bar graphs and line graphs. Then, I chose the best representative graphs for my data that are the easiest visually. After choosing the graphs, I started coding the graphs using python. In order to visualize them on my Streamlit WebApp.

I wanted to check the following from my data:

- The first thing that I visualized was the total number of jobs that I'm going to analyze, which are 923 jobs.

Image 9

Number of Scrapped Jobs

- The below graph shows the months the jobs were scrapped. These months are May, June, July and August. The graph shows that the most jobs that were scrapped were posted on August.

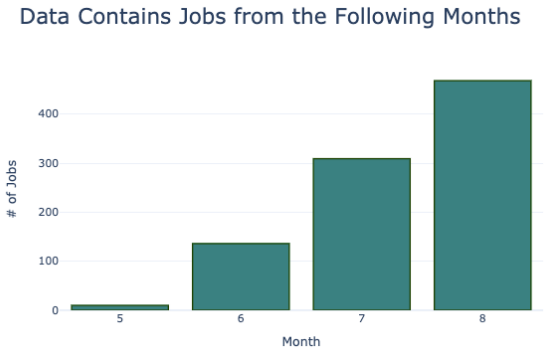


Image 10

Data scrapped is months May, June, July and Ausgust

- The below graph shows location of job openings along with city names. It shows that Beirut is the city that has the most job openings. It has 77% of job openings. The second city with most job openings is Mount Lebanon (15%), followed by North Lebanon (3.4%), South Lebanon (3.3%), Beqaa (0.6%), and Akkar (0.4%).

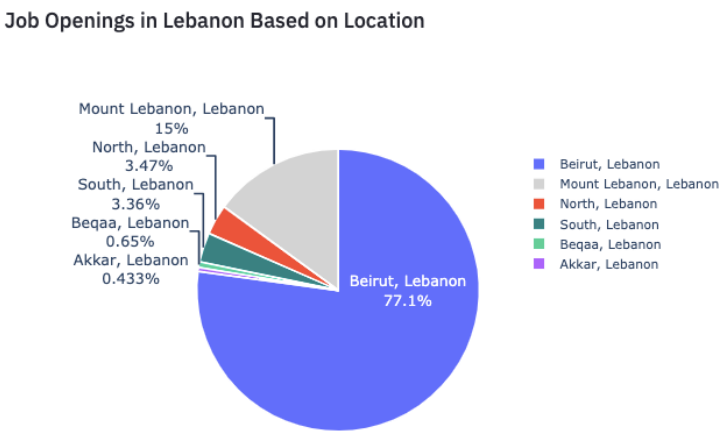


Image 11

Job openings based on Location

- The below graph shows the most demanded job openings in Lebanon based on the LinkedIn data that I scrapped. It shows that the most job opening in Lebanon is "Manager" followed by "Software Developer", "Sales Person", "Designer", "Engineer", "Analyst", and "Marketer"…

Image 12

Most demanded job positions in Lebanon

- The below graph demonstrates the most demanded job openings based on cities in Lebanon. You can filter based on different cities in Lebanon. For example, in South Lebanon, the most demanded job position is “Software Developer”, followed by “Manager” and “Teacher”.

Image 13

Most demanded jobs based on cities in Lebanon

- The blow graph shows companies in Lebanon that have the most job openings. As we can see, the company that has the most job openings is “Toptal”, followed by “Agoda”, “Monty Mobile”, “Crossover for work”, “Presential”, “Novelus”, and “Pepsico”.

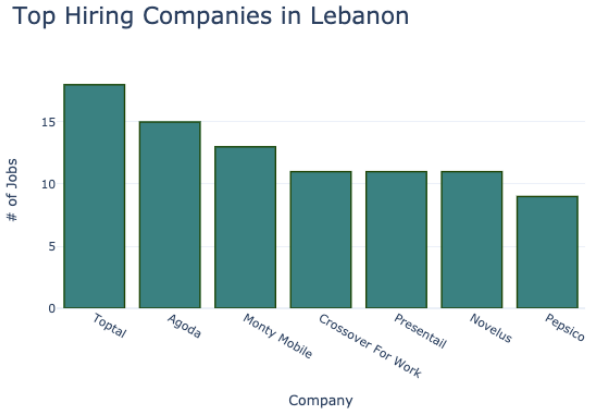


Image 14

Top hiring companies in Lebanon

For the second part of my project

In this section, I found some difficulty in creating the graphs while filtering based on professions. However, with some trial and error I was able to create the tables and graphs correctly.

- In the below table, on my Streamlit WebApp, you can filter based on the profession you want to know more about and a table will be shown of the specific tasks for the profession you chose.

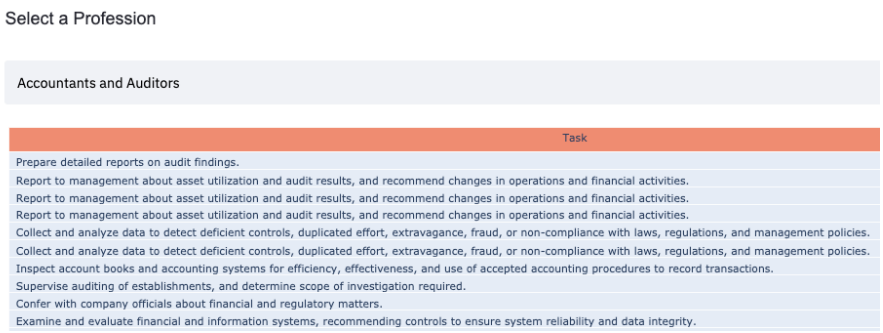


Image 15

Tasks for different professions

- In the graph below, on my Streamlit WebApp, you can filter based on the profession you want to know more about and a bar graph will be shown of the specific skills for the profession you chose, and the skills will be shown based on their importance.

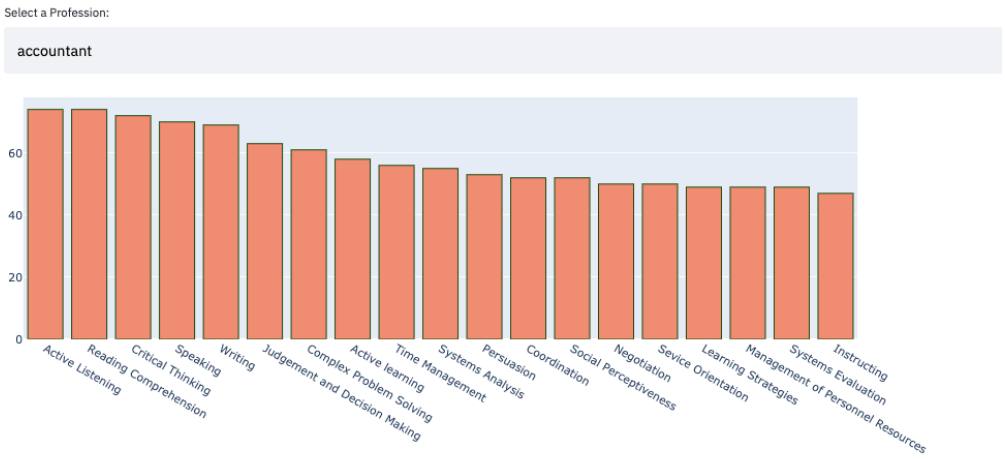


Image 16

Skills for different professions

- In the graph below, on my Streamlit WebApp, you can filter based on the profession you want to know more about and a table will be shown of the detailed work activities for the profession you chose.

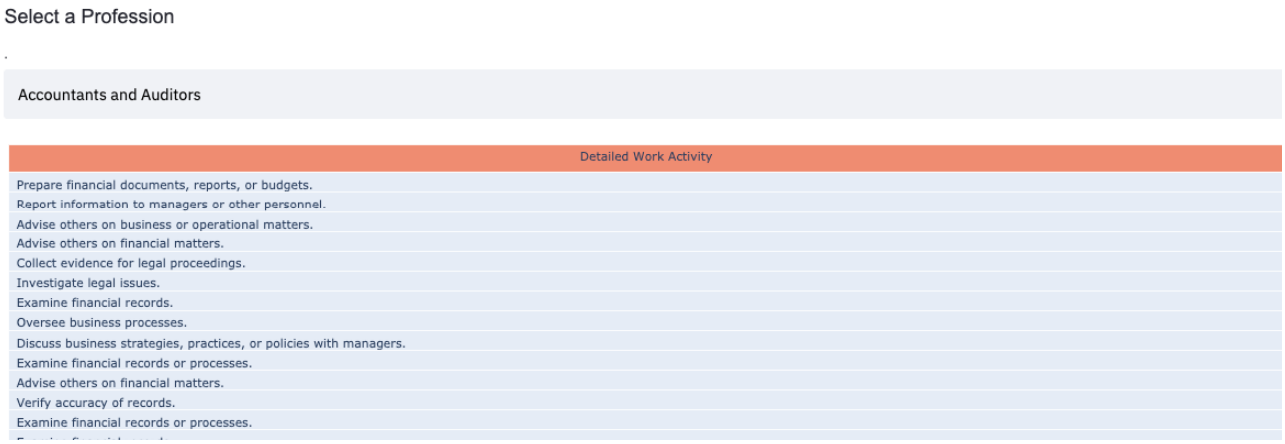


Image 17

Detailed work activities for different professions

5. Discussion

The key findings that I was able to interpret from my data could help Lebanese students better choose their career choice. The key findings from my project are that the most demanded job in Lebanon is “Manager”, “Software Developer”, and “Sales Person”. From here Lebanese students might be more motivated to do majors that are related to these job positions so that they could find jobs for their skills upon graduation. Also, students who live in different cities could check the most demanded jobs in their cities and not only in Lebanon. In this project, the top hiring companies are also mentioned. Therefore, students could think of applying to the most hiring companies if their skills match for an internship. This could help students get an offer for a job later on, since the company is willing to hire. Furthermore, having working labor with the rights skills that fills the country’s needs boosts productivity and hence the economy.

Saying you love to be a chemical engineer, but not knowing what are the everyday daily job of a chemical engineer is a problem. Many people end up in the wrong majors because they didn’t know what are the daily tasks for their chosen profession. That’s why going through a detailed description of the profession you choose is very important. I built a detailed description that contains the tasks, skills and detailed work activities for almost every profession in the second part of my project – the part about the job information. In this part not only, Lebanese students could benefit from it but also all other students in the world who want to get more detailed information about a certain profession. This part helps students find the profession they love to do.

Many things could be added to this capstone project for enhancement. Scraping LinkedIn skills and scoring them would be an added value. Lebanese students would be able to know the skills that

companies need in Lebanon. Also, what could be improved upon this project is if the Streamlit WebApp that I created directly scrapes data from LinkedIn on the spot. However, LinkedIn puts some restrictions on scrapping the data directly. In addition, having a larger dataset would result in much more accuracy for my analysis.

6. Conclusion

"Career Dashboard Guide for Students in Lebanon", the capstone project that I built for my final year project for my master's in Business Analytics at the American University of Beirut is a simple online tool available to all students to explore job market trends in Lebanon and guide their future, helping them explore different kind of professions, make decisions about their future, and prepare them for the next step in their education and career planning. I scrapped data from LinkedIn to get information about job positions in Lebanon and I scrapped data from O*net website to get information about job professions. After scrapping and transforming the data into CSV files, I cleaned it, preprocessed it and added necessary columns for further analysis. Next, I visualized my data using python on Streamlit WebApp so that it would be accessible by any student online. The WebApp shows most demanded jobs, most hiring companies, cities with the most job openings all in Lebanon. It also shows information about tasks, skills and detailed work activities of different job professions. This helps the problem that students in Lebanon are facing, which is that they don't know what are the demanded jobs in Lebanon so that upon graduation they would be able to find jobs.

Some future suggestions for my project, I suggest scrapping data not only from LinkedIn but also from other sources online and to ask “Ministry of Labor” if they have any datasets available about job openings in Lebanon. This would make my analysis much more accurate. Also, I suggest scrapping more information about the job opening because it would add more benefit to my project. Such as, scrapping industry, skills, job type and experience level. Also, adding another page that asks questions about personality, hobbies and skills and predicts or recommends the profession that best suits the student. In addition, scrapping skills from the job openings in Lebanon would build a more accurate picture for students to know what are the exact skills needed for certain professions in Lebanon. These suggestions would add more usefulness to my capstone project, you can find the python notebook and the Streamlit app link on my GitHub account @lunabaalbaki. I hope my project would help as much students as possible.

7. References

- Chaudhary, D., Prajapati, H., Rathod, R., Patel, P., & Gurjwar, R. K. (2019). Student Future Prediction Using Machine Learning. https://www.mendeley.com/catalogue/1aa26861-e33b-3af3-a570-9735fb98d04d/
- Najwa Yaacoub, (2011), The Labor Market in Lebanon. https://daleel-madani.org/sites/default/files/Resources/CAS_Labour_Market_In_Lebanon_SIF1_0.pdf
- Claudia Knobloch, (2018), German Cooperation, Employment and Labor Market Analysis in Lebanon. https://www.giz.de/en/downloads/ELMA_Lebanon_2019.pdf
- Statistics in Focus, (2011), The Labor Market in Lebanon, https://daleel-madani.org/ar/civil-society-directory/central-administration-statistics/resources/statistics-focus-labour-market-lebanon
- ILO, (2017), Labour Market Information Review and Analysis, https://www.ilo.org/wcmsp5/groups/public/---arabstates/---ro-beirut/documents/publication/wcms_559670.pdf