

Name: Sagar Khode

Email address: sagar.khode1@gmail.com

Contact number: 7020624724

Anydesk address:

Years of Work Experience: Final year engineering student(final semester)

Date: 03-06-2021

Self Case Study -1: * Career Village Recommendation Engine *****

“After you have completed the document, please submit it in the classroom in the pdf format.”

Please check this video before you get started:

https://www.youtube.com/watch?time_continue=1&v=LBGU1_JO3kg

Overview

*** Write an overview of the case study that you are working on. (**MINIMUM 200 words**) ***

- CareerVillage is a non-profit organization that helps high school students, unemployed undergraduates, one who is need of work to get career advice online from real-world professionals. Professionals can then answer those questions from any laptop or smartphone, giving students the inspiration and education they need to make better decisions about their futures.
- Issue :- Please refer research paper remarks .Any one can understand what happen happens when there is no proper guidance at proper time.
- CareerVillage is a motivation to help me for this case study. This case does not belongs to a single person, community, or a organisation.

- Its a complete path way to one who wants career guidance , our aim is to build robust technique so that as someone needs help in knowledge ,information then from professionals who have huge experience in particular domain or field will help to that particular person
- For build this case study we are using datasets which are provided by CareerVillage.
- As we know youth and professional are the important part of Society , because if one learn he/she not only helps to his/her family but also complete humanity.
- I want to ellaborate name behind the organisation CareerVillage

Village → where lots of people stays with peacefull nature , with helping each other ,with taking care of each-other . In result long journey with less trouble in life(think as you are in village.)

Career →Career is combination of educational knowledge , proper information and best mentally support(given by mentor[any professional who have long work experience])

Career = educational knowledge + proper information + best mentally support

Research-Papers/Solutions/Architectures/Kernels

*** Mention the urls of existing research-papers/solutions/kernels on your problem statement and in your own words write a detailed summary for each one of them. If needed you can include images or explain with your own diagrams. **it is mandatory to write a brief description about that paper. Without understanding of the resource please don't mention it*****

Sr No.	Paper Name	Publication	Date of submit & paper Number	Abstract	Remark
01	THE FACTORS AFFECTING CAREER GUIDANCE AND CURRENT STATUS OF CAREER GUIDANCE	PROBLEMS OF EDUCATION IN THE 21stCENTURY	Volume 48, 2012 ISSN 1822-7864	The purpose of the present study is to investigate the factors affecting career guidance	<ul style="list-style-type: none"> • US started career guidance in 1909 , and Career Office by Frank Parson in Boston(Father of career Guidance). Career models were started be called as

<p>SERVICES IN TURKEY</p>	<p>and their reflections of the career guidance services given in the country. Based on a literature review, it is known that career guidance services can be a part of economic and social development. In the present study, these factors are analyzed under these headings as the structure of the region, economic conditions (labor force of the country and employment), historical and socio-cultural factors (family structure, changing values) and educational factors. Moreover, it is emphasized that the country is the only secular country and only candidate for</p>	<p>developmental career guidance within the framework of vocational guidance. As time goes on from 1909, 1951, 1960, 1970, 1996 there is increase in importance on Career guidance. Career guidance and career counseling are both old and new terms (Herr and Cramer, 1996). They also stated (1996) that it has a history of 75 years, that is, it has a history of 90 years up to now.</p> <ul style="list-style-type: none"> • Schools provides such guidance to students who are on the point of selecting a profession or going on studying at higher levels; public employment agencies; on the other hand, help the unemployed looking for a job (OECD, 2004). • The Turkish history has three main axes. These are traditions brought from Middle Asia to Anatolia, Islamic period and Imperial era (Güngör, 1986, 107) • The Complete paper is based on Growth of Turkish Educational
---------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

			<p>the European Union membership among the Muslim communities. After drawing a framework with these factors, the state of the country in guidance services and its reflections abroad were investigated. It is of great importance to draw the attention to such issues in Turkey and to develop "career guidance programs" as in developed countries.</p>	<p>Process.</p> <ul style="list-style-type: none"> They want to explain what was the previous condition of Turkey before Republic Democratic's , How people started working and started focusing on Career Guidance. In the begining they were unable to focus because lack of knowledge. But as the time goes the start developing there countries by helping Youth by Career guidance.
02	The Effects of Career Choice Guidance on Identity Development	Hindawi Publishing Corporation Education Research International	<p>Volume 2013, Article ID 901718 http://dx.doi.org/10.1155/2013/901718</p>	<p>The effects of a career guidance program on the identity development of students with career choice problems. Starting from the perspective of identity theory, we argue that career choice intervention in late adolescents</p> <ul style="list-style-type: none"> Career choice is seen as one of the major tasks development. In Erikson's life-span stage theory, identity development is the main developmental task in late adolescence life cycle. Commitment formation means making choices about what really matters to them, and in which they are personally involved.

should not only address the commitment formation in the vocational domain, but also in the personal and global domain

- School and career are the most important domains in which commitments have to be chosen said in this article.
- Here, they focuses on commitemtent which they were saying commitemtent, to the career and for the career.
- United States started going to school,colleges,uni versities to give career guidance to many of students ,so that they can have good career in future.
- Groningen Identity Development Scale (GIDS) [3], is a semistructured interview that measures the level of exploration and of commitment strength and the commitment content in different domains. It focuses on six domains: philosophy of life, parents, friendship, vocational domain, personal domain and intimate relationships, and global identity
- This was the innovative idea to go and guide youth about there career.

03

CAREER GUIDANCE CHALLENGES FOR THE NEW CENTURY UNDER AN INTERNATIONAL PERSPECTIVE	Orientación y Sociedad – 2003/2004 – Vol. 4	Orientación y Sociedad – 2003/2004 Vol. 4 UNLP, La Plata, 24 – 26 September 2003.	This paper is dedicated to key issues of the actual challenges in all societies regardless their development al level and how the international guidance community is coping with these challenges. It deals with the importance of guidance in a changing society, quality assurance, access to services and qualification of guidance practitioners under an international perspective.	<ul style="list-style-type: none">• This paper was wrote to encourage the efforts of all professionals in country to promote the advancement of career guidance in a time of serious social problems.• Dr. Bernhard Jenschke said ,Globalisation at the beginning of the new millennium and its impacts on human and social life there is an urgent need to rethink guidance and counselling.• And to overcome all the past problems , one should know what is the proper use of good's and resources ,for which guiding future fellow is must important thing to achieve.• Dr. Bernhard Jenschke also said Lifelong learning needs lifelong guidance.• He also discussed about a book on “ Technical and Vocational Education and Training in the Twenty -first Century: New Roles and Challenges for Guidance and Counselling” (UNESCO/IAEVG/I
----------------------------------------------------------------------------------------------------------------------	---------------------------------------------------	-------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

AC, 2002) where several international experts focus on the crucial role of guidance to enhance the development of technical and vocational education and training.

- They were focusing on key tool, i.e to achieve these goals a new thinking of guidance and counselling has to ensure that everyone can easily access good quality information about learning opportunities and personal advise how to combine living and working and to pursue as self-motivated and active citizens.
- In Short this article was telling professionals to focus on guidance to newbies. So that they can survive, make to know how to use those valuable resources and skills to acquire new generation obstacle.
- Here the used method by taking help of conferences, books and public debates.

04

Career Guidance in Secondary Schools - A Literature Review and Strategic Solutions for Vietnamese Rural Areas	American International Journal of Social Science	Vol. 4, No. 5; October 2015 ISSN 2325-4149	This paper reviews the role of career guidance and counseling at schools across the globe. The aim of this paper is to overview the career guidance in schools through literature review in order to bring this field into perspective. Furthermore, it attempts to find orientation and solutions in rural Vietnamese contexts. With the socio-economic development in rural areas and increasing rural to urban migration, it is necessary to enhance the quality of education in career value, based on the students' interests, abilities and passion, with more students	<ul style="list-style-type: none">• Vietnamese labor resources stated that "excess of higher education degrees but lack of skilled workers". It means workers(person/individual) have good education but lack of skills because of lack of guidance in proper way to goal.• This paper says in 9.2014, there were 1,74,000 undergraduates in Vietnam. This is a serious issue because if you cant earn or have your goal then what is the use of this high demand education.• If the teachers , parents or mentors dont help in guidance to a learner then how would he know what to do about his/her career.• After seeing worst conditions of unemployed human , many countries started giving guidance to everyone from pre school to post education where ministry taken the complete charge for this serious issue.• Not only the countries but also continents by as much as possible
---------------------------------------------------------------------------------------------------------------	--------------------------------------------------	-----------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

05	<p>Career Guidance, Participation of Students and its Implication for Kano, Nigeria</p>	<p>The Malaysian Online Journal of Educational Science</p>	<p>Volume 1, Issue 3 [1] Institute of Education International Islamic University Malaysia isabakar2001@yahoo.co.uk</p>	<p>switching to vocational education for their future career opportunities</p>	<p>means of methods.</p> <ul style="list-style-type: none"> After which trend of Rebuilding of Rural Areas in Vietnam in the period 2010-2015 and Directional Solution for Career Guidance started.
				<p>The purpose of the study was to examine students' participation in career guidance activities as mechanism through which young adults are developed into productive, responsible personalities well equipped for life and work in today's technology based society. The study focused on career information search, career exploration and assessment aspects of participation.</p>	<ul style="list-style-type: none"> Rosemary (2002): says "acquiring knowledge self knowledge, developing specific career and educational goal, adjusting to changing conditions, planning career and educational programme to achieve goals, developing problem-solving and decision making skills, coping with the outcome of decisions, and enhancing social, emotional and cognitive skills" Career guidance explains , It involves the person's creation of a career pattern, decision-making style, integration of life roles, values expression, and life-role self concepts (Herr & Cramer, 1996)[Father of Career Guidance]. Holland's theory says choosing a career, means

				people search for environments that will let them use their skills and abilities, and express their attitudes and values.
06	The impact of career guidance on progression in learning and work: a literature review	Hooley, T. et al (2016) 'The impact of career guidance on progression in learning and work: a literature review' Derby: University of Derby	Derby: University of Derby. ISBN: 978-1-910755-10-5	<p>This article explain the 1st use of machine learning in our Survey, here they used statistics, calculated Mean, Standard Deviation, t-Values for Gender Participation in Career Guidance Activities' Subscale. And they also applied regression analysis on what data they collected. But as per my concern there may not be good result after all the comparisions</p> <ul style="list-style-type: none"> • This paper is amazing and have very important two points <ol style="list-style-type: none"> 1. Services need to be provided in a timely fashion, and as quickly after an individual has dropped out of learning or work as possible. 2. Services need to be provided professionally by skilled advisers <p>This 2 points tells everything about this paper intension.</p> • This paper tells about 5 important questions to focused by National

	<p>career guidance, although access to career guidance is only one amongst a range of factors that might influence an individual's likelihood of progressing. It also notes that progression can be difficult to measure in research. The initial review found a range of evidence which demonstrate that career guidance can have a positive influence on individuals' progression to learning and work. It highlighted a number of features that underpin the effectiveness of career guidance in this area.</p>	Careers Service
	<ul style="list-style-type: none"> • 1. What evidence exists which describes the policies, systems and processes that underpin the organisation of national careers services? • 2. How can careers services maximise customer satisfaction? • 3. How can careers services maximise their impact on career management skills? • 4. How can careers services maximise their impact on individual's progression to positive learning and work destinations? • 5. How is effective brokerage between education and employers organised? • According to OECD, Career guidance is defined as "services and activities intended to assist individuals, of any age and at any point throughout their lives, to make educational, training and occupational choices and to manage their 	

careers" (OECD, 2004).

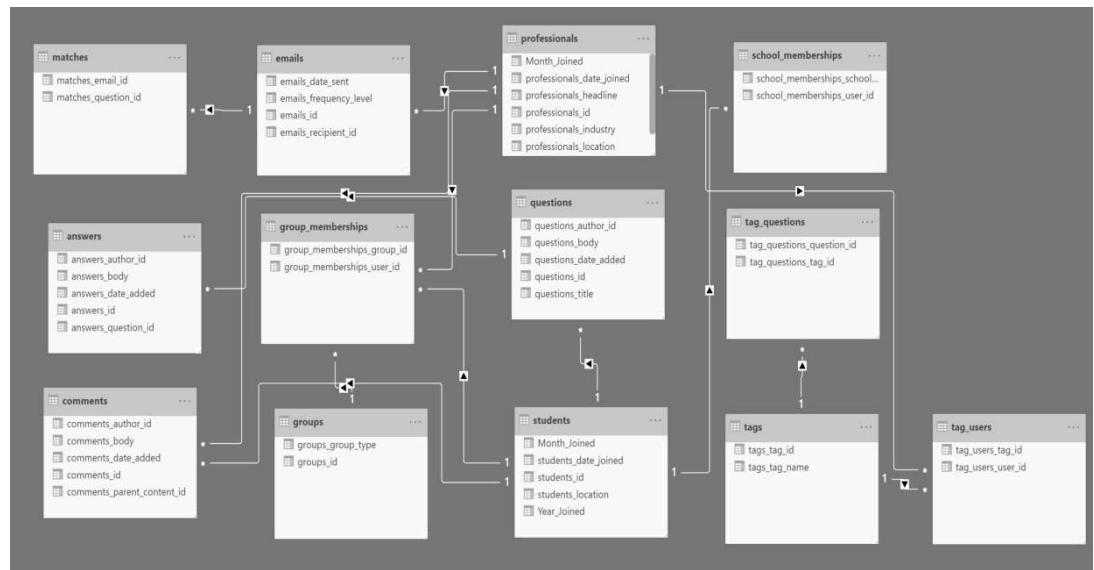
- This paper explain support is very import to get proper career flow, support in the sense in professionalism.
- To get proper path or career guidance one should have proper professional behaviour or connection in any network like social media , relatives, personal friend circle and so on...
- To be in good life 1st have good connection said in various researches.

First Cut Approach

*** Explain in steps about how you want to approach this problem and the initial experiments that you want to do. (*MINIMUM 200 words*) ***

*** When you are doing the basic EDA and building the First Cut Approach you should not refer any blogs or papers ***

- Aim :- Our aim is to build a model which can help to everyone(this is not specific to high-school students but also to school, undergraduate, graduate and also some limit of professionals too.)
- Our Dataset consist of following file's
 - ['answers.csv', 'answer_scores.csv', 'Case Study 1.ipynb', 'comments.csv', 'DATA', 'emails.csv', 'groups.csv', 'group_memberships.csv', 'matches.csv', 'professionals.csv', 'questions.csv', 'question_scores.csv', 'school_memberships.csv', 'students.csv', 'tags.csv', 'tag_questions.csv', 'tag_users.csv']
- Relation Between them are as follows:-



- Our task here is “when some students or learner ask or post question related to career guidance (questions like
- 1). “I have good knowledge of math but because of low knowledge of real world what should I do ? ”
- 2). “I have completed my B.com in 2020, and want to be in HR domain . what should I do further ? ”
- 3). “Want to study in Abroad , but because of less money and middle class family . What should I do ? ” etc etc
- For experience person this questions are like silly easy problem right !!!!

- So, here professionals who have very much experience in particular or respective domain will help those learner's or beginner's.
- Exact first cut approach starts from here
 - ❖ By using machine learning with the help of python we fetch that collected data to different variable's.
 - ❖ Performing EDA :- we will calculate professional's detail as shown below...

Professionals Detail :-

Total Professionals: 28152

Industries they belong to: 2470

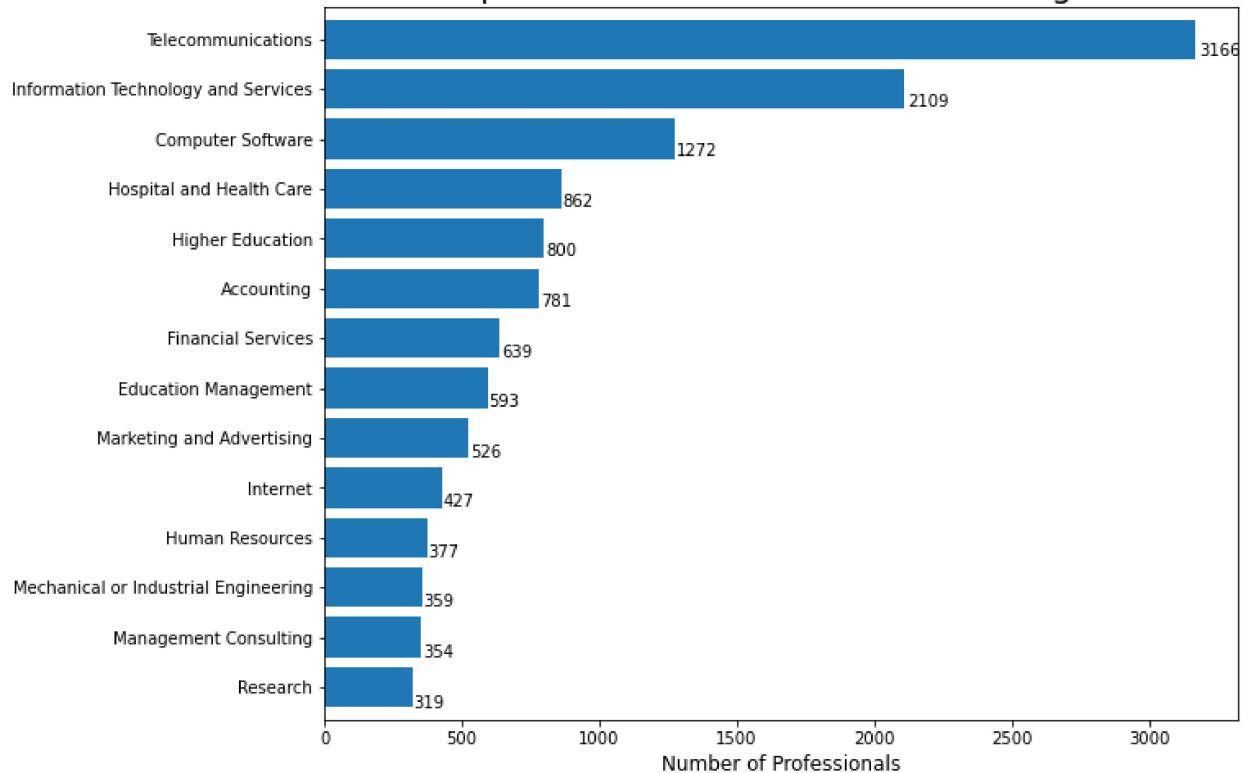
Tagged to Unknown Industries: 2576

Locations they belong to: 2582

Tagged to Unknown Locations: 3098

- ❖ Now we will find top 14 industries professional belongs to

Top 14 industries Professionals belong to



As you Can see, Most of the professionals are from Telecommunication , IT Services , Computer Software field.

In 1st position Telecommunication with 3166 Professionals.

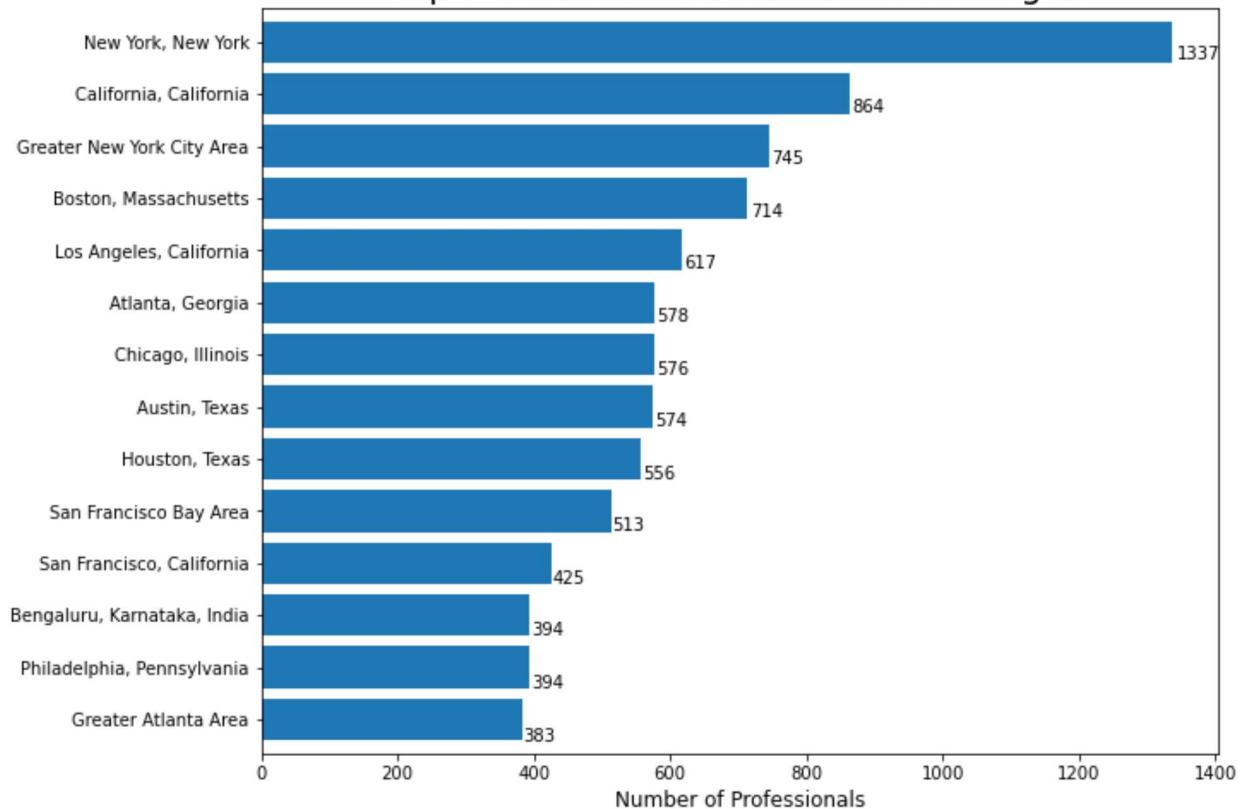
2nd position IT Services with 2109 Professionals.

3rd position Computer Software field with 1272 Professionals.

And so on...

- ❖ We find out top 14 location's professional's belong to

Top 14 Locations Professionals belong to



As you Can see, Most of the professionals Locates are from New York , California , Greater New York City Area.

In 1st position New York with 1337 Professionals.

2nd position California with 864 Professionals.

3rd position Greater New York City Area with 745 Professionals.

And so on...

- ❖ Then we got how many professional's joined year wise

Count

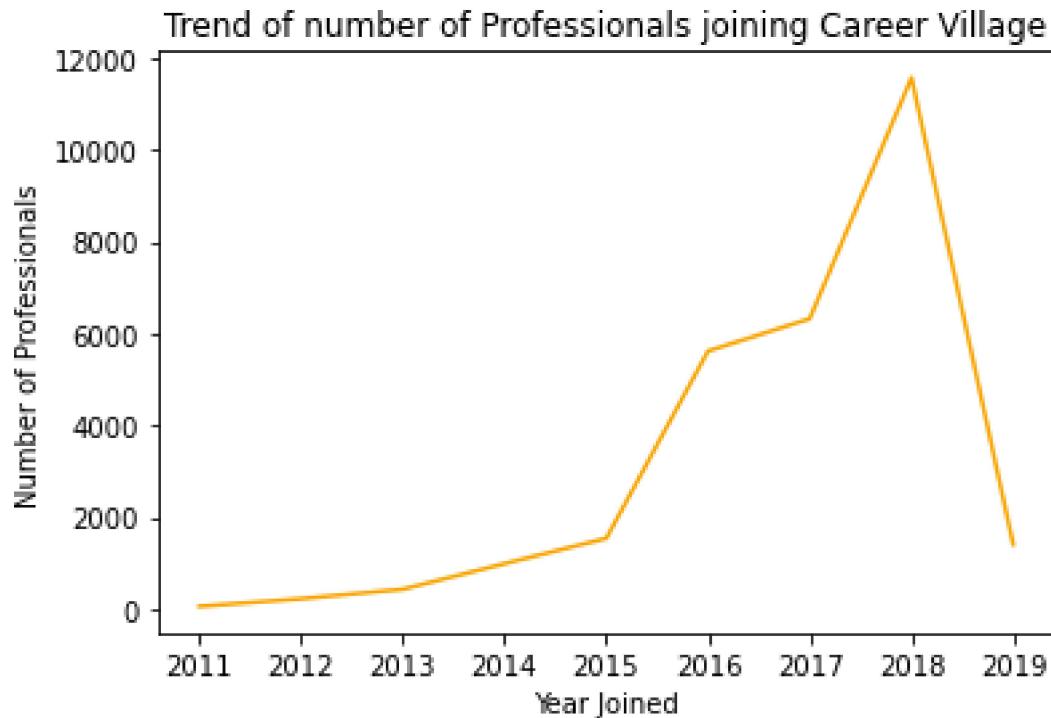
YearJoined

2011 57

Count	
YearJoined	Count
2012	222
2013	429
2014	987
2015	1540

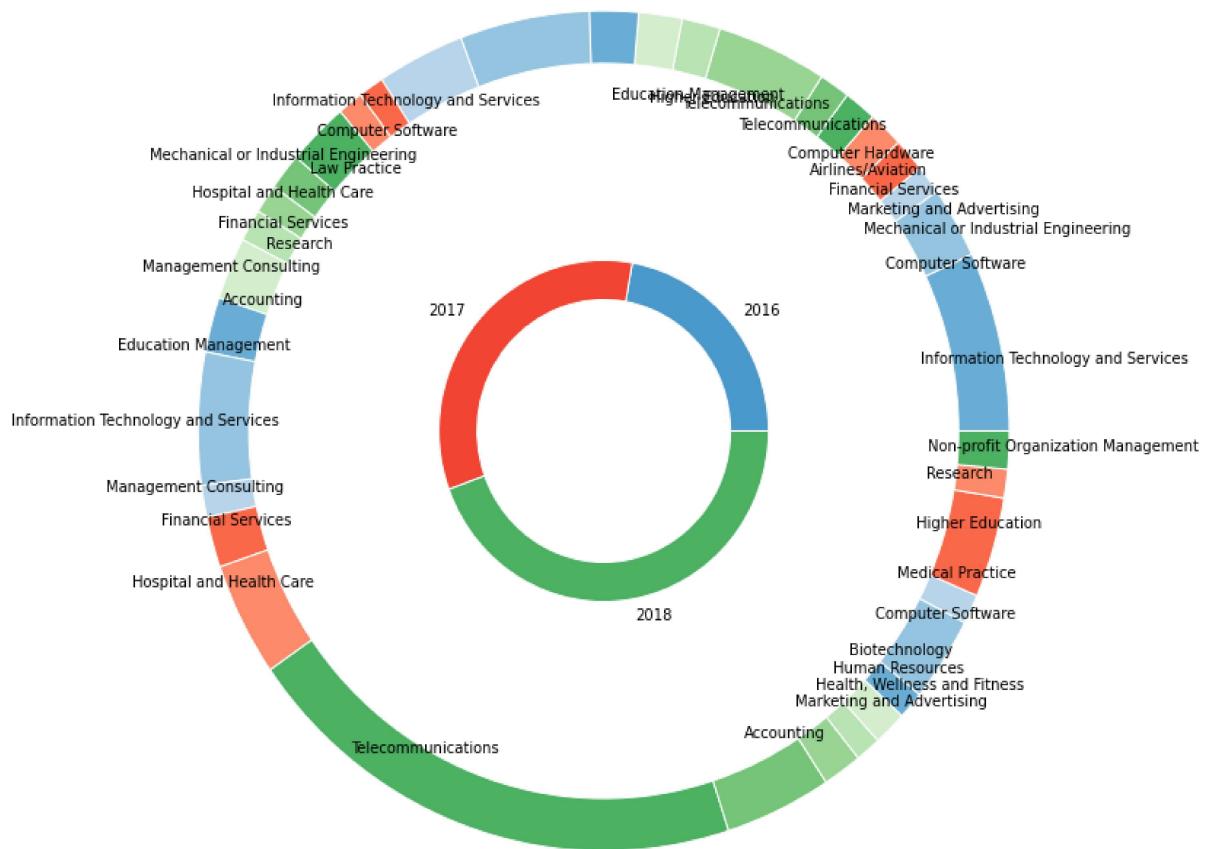
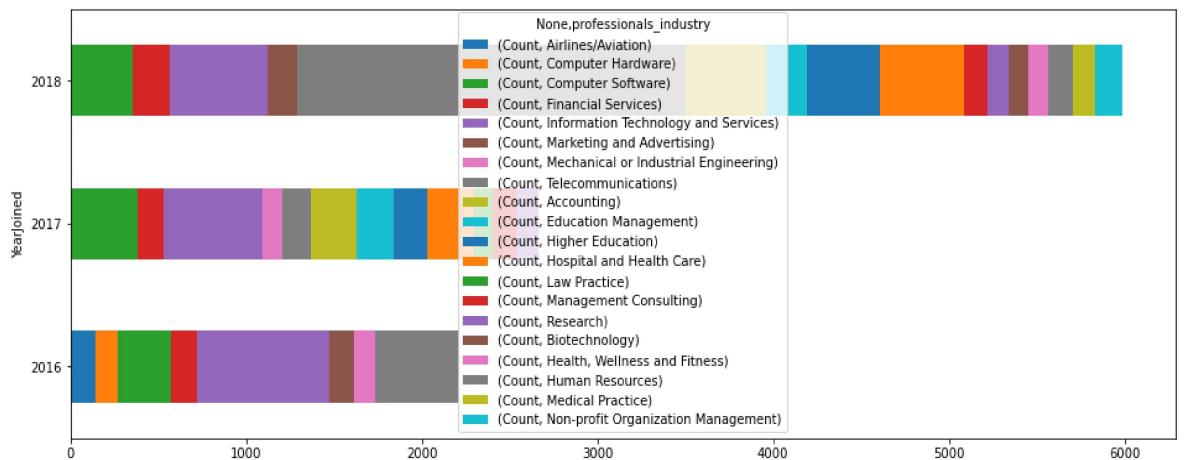
Date of Joining(in Year) of Professionals with Counts

- ❖ We done this because we must know what is the progress of professionals and also the count.
- ❖ Now we will show it graphically how much supporter professionals are there in CareerVillage.



By Observations , Since majority have joined from year 2016 to year 2018, it would be good to know the industry wise count in these years

❖ So , see 2016-2018 trend of professional's



Observation:

Telecommunications in 2018

IT Services in 2016

and Hospital and Health Care in 2017 were the more to joined Career Village

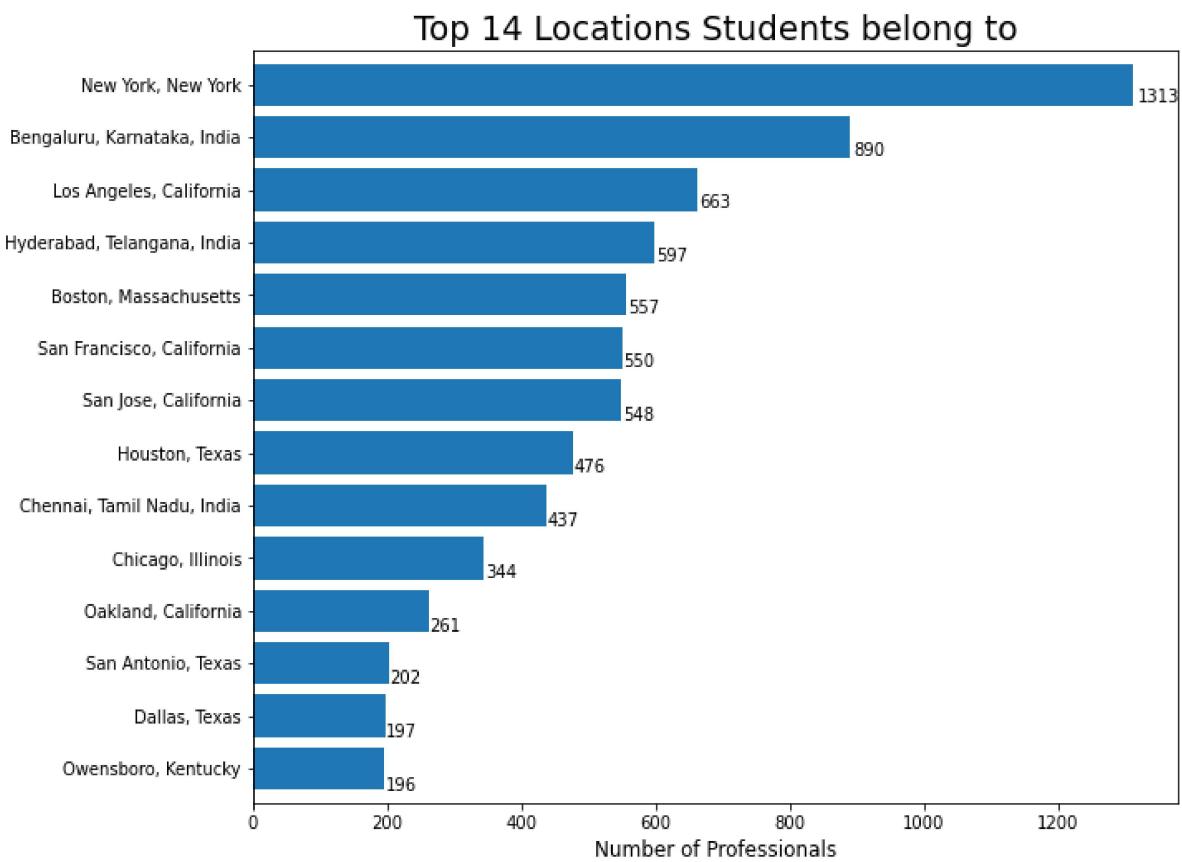
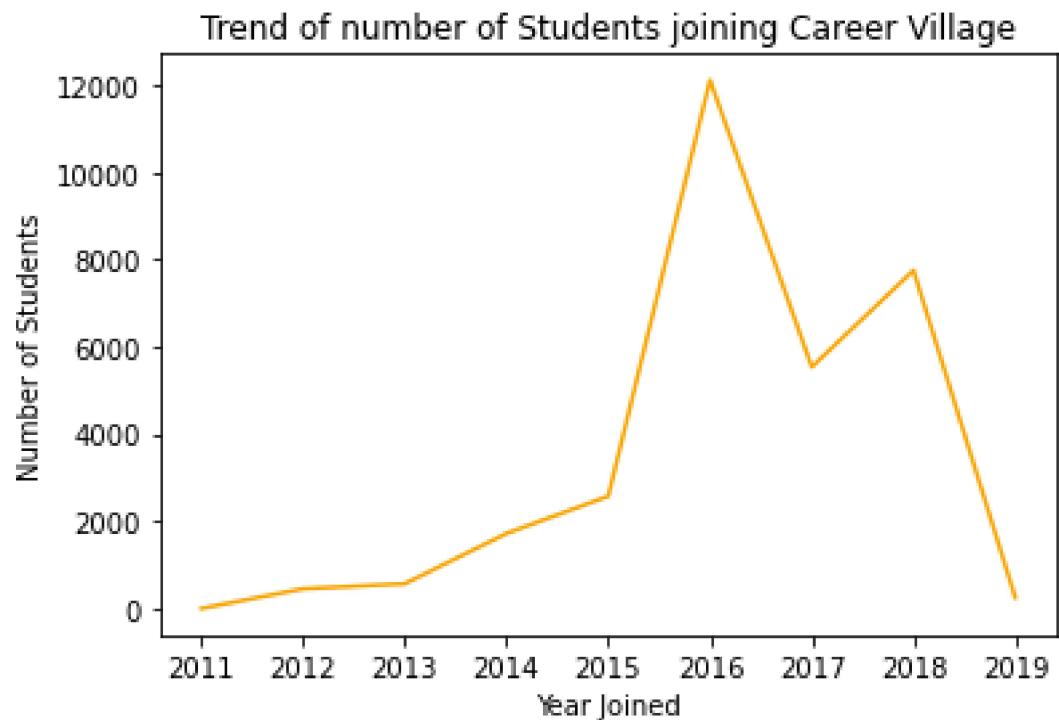
- ❖ Now, as above we had calculated Total no. of students, Location's they belongs to and tag to unknown location's. And also the count of student's year wise , showed in plot also,

Total Students: 30,971

Locations they belong to: 5,480

Tagged to Unknown Locations: 2,033

Count	
YearJoined	
2011	2
2012	453
2013	570
2014	1722
2015	2576



❖ Calculated Groups of group i.e

groups_group_type	Count
cause	2
club	1
competition	1
interest group	1
mentorship program	3
professional network	8
youth program	33
Total Groups :-	49

- ❖ Now calculated Total Groups of Student's and Total Groups of Professional's

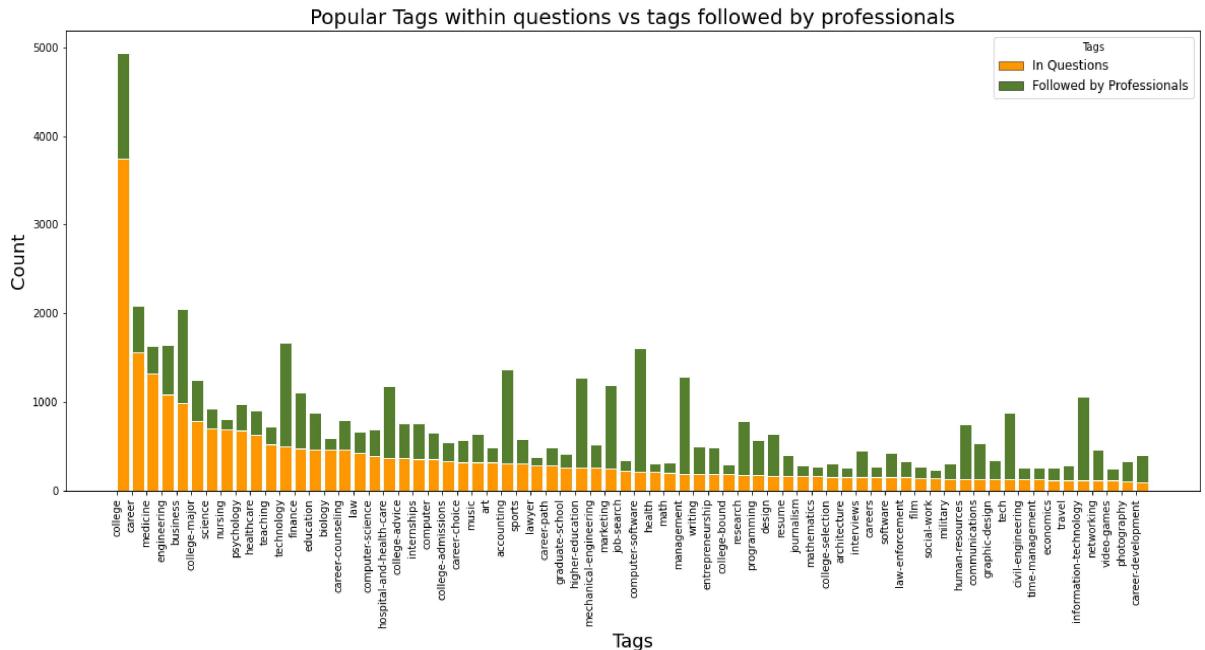
	Professionals_Count	Students_Count
groups_group_type		
cause	161	71
club	12	10
competition	3	1
interest group	9	4
mentorship program	17	5
professional network	64	28
youth program	461	192
Professionals_Count	727	
Students_Count	311	

- We will move 2nd part i.e Question's and Answer's part

- ❖ Now , we calculated Total number of question and tags

Number of total tags are 176350 for 49437 questions

- ❖ We , combined questions and tags.
- ❖ Tags which are within the question and the tags which are given by professionals.
- ❖ Also showed in stacked plot

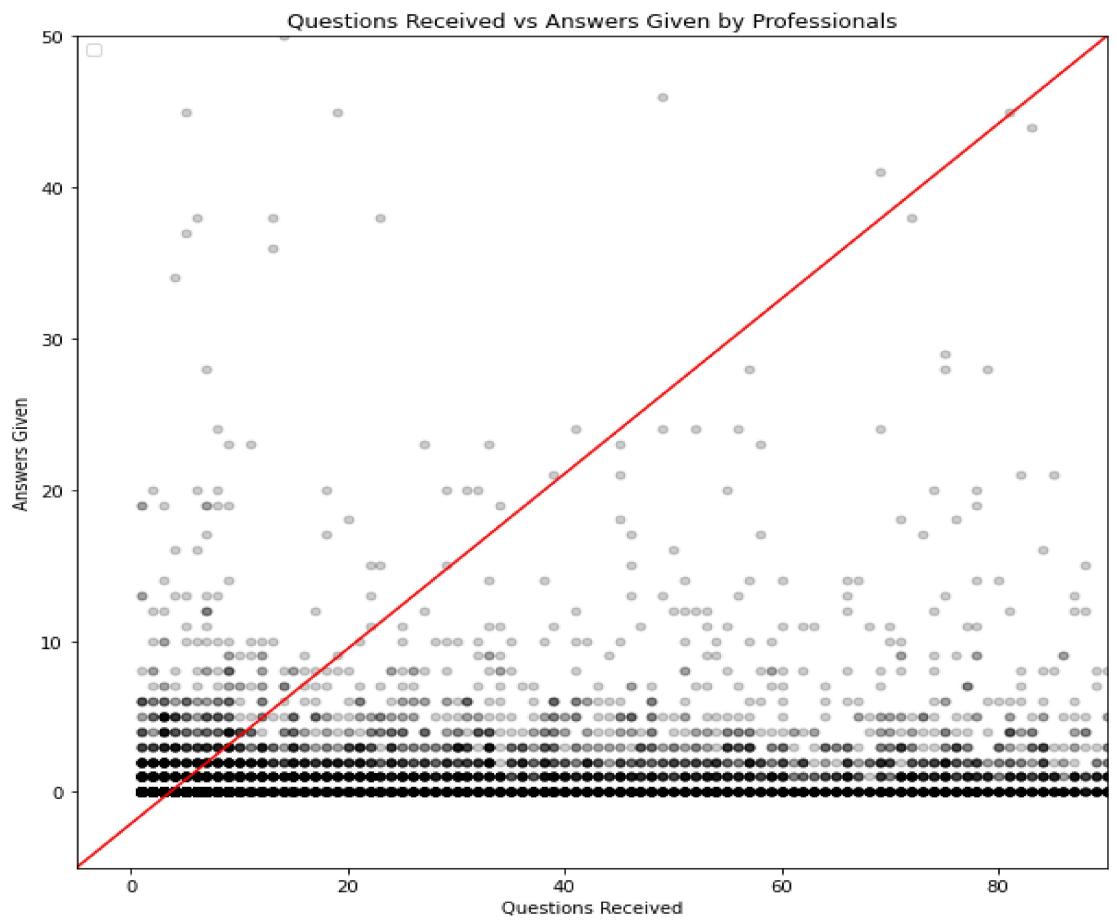


❖ Showed summary of Professionals detail .

	emails_recipient_id	professionals_date_joined	professionals_location	professionals_industry	emails_received	questions_received	answers_given
0	00009a0f9bda43eba4710 4e9ac62aff5	2016-03-14 17:00:48+00:00	New York, New York	Media	14	15	3.0
1	000196ef8db54b9a86ae7 0ad31745d04	2018-05-15 20:27:44+00:00	Greater Chicago Area	Accounting	42	109	0.0
2	000d4635e5da41e3bfd83 677ee11ddaa4	2016-04-27 16:38:34+00:00	Nashville, Tennessee	Information Technology and Services	945	2335	3.0
3	000e2b5714444d79a672 bf927905135c	2017-10-11 02:04:27+00:00	Greater Detroit Area	Financial Services	3	6	0.0
4	0018873fbf7742aba1bf13 fff12cbfa4	2018-05-24 20:34:27+00:00	Portland, Oregon Area	Financial Services	10	25	0.0

This are top 5 rows only.

❖ Now will show email got and answered by professional's



Observation :-

Scatterplot for Questions Received vs Answers given

Shows that Majority of them have been sent large number of questions but response is low

Whereas some professionals have shown interest in answering more than questions being mailed

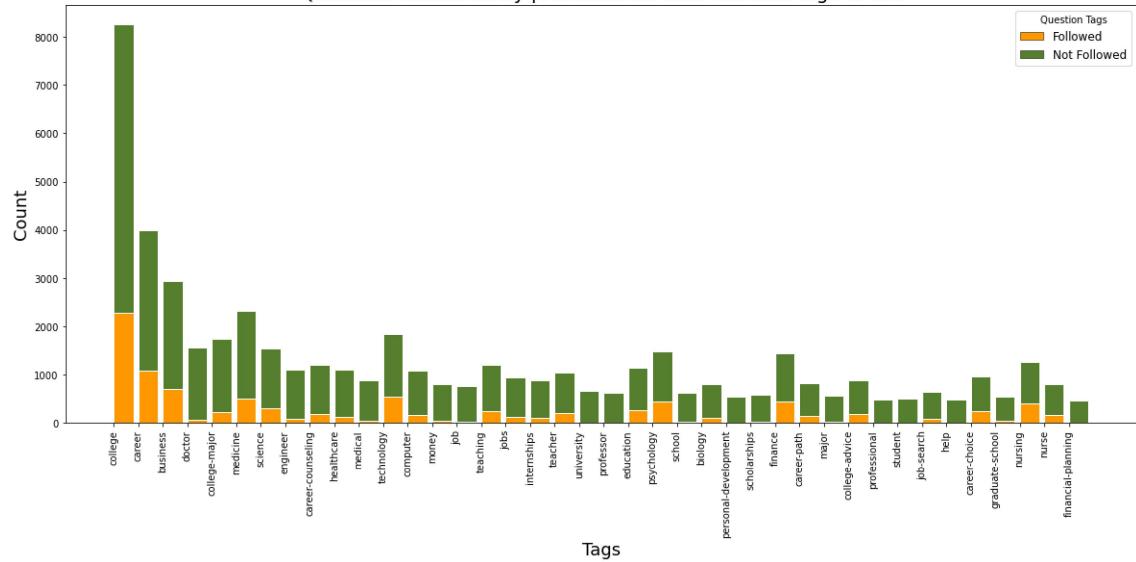
❖ Now , merging all tags as shown below

	tag_questions_tag_id	tag_users_tag_id	Count	tags_tag_name
0	27	-1.0	161	military
1	27	27.0	142	military

	tag_questions_tag_id	tag_users_tag_id	Count	tags_tag_name
2	29	-1.0	949	teaching
3	29	29.0	253	teaching
4	36	-1.0	376	health

❖ Shown , Questions answered by professionals who followed tags or not

Questions answered by professionals who followed tags or not



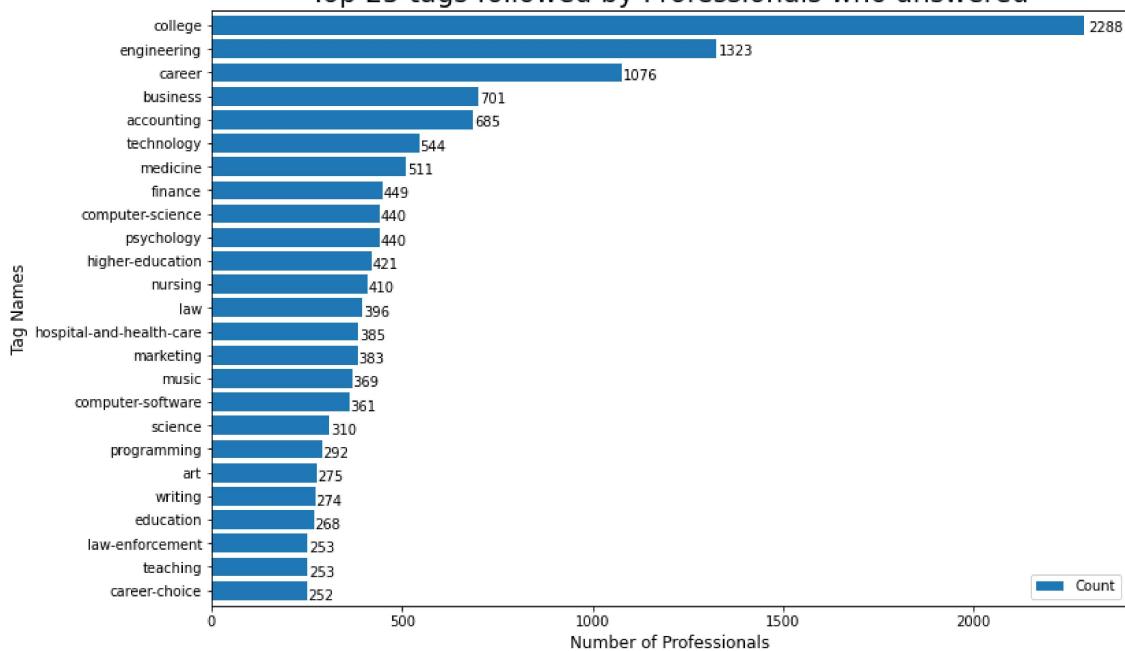
Stacked Bar chart to show top 40 tags

whether professionals followed it or not when they answered questions

This shows that there are majority of answers given by professionals without mail notification

❖ Top 25 tags followed by Professionals who answered

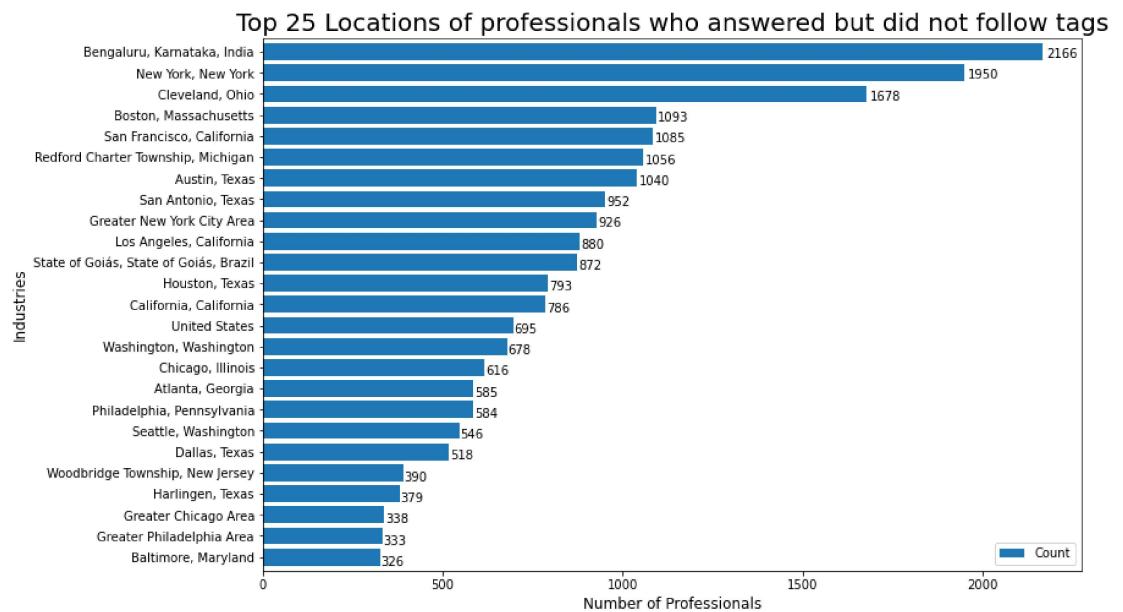
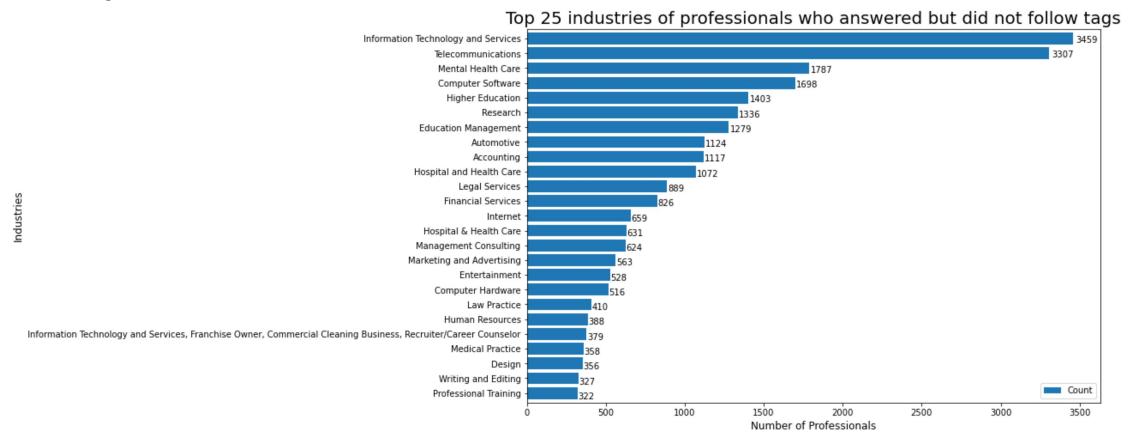
Top 25 tags followed by Professionals who answered



❖ Created table with professionals with joining date and time.

Count		professionals_id	professionals_location	professionals_industry	professionals_date_joined
0	3	00009a0f9bda43eba47104e9ac62aff5	New York, New York	Media	2016-03-14 17:00:48+00:00
1	3	000d4635e5da41e3bfd83677ee11dda4	Nashville, Tennessee	Information Technology and Services	2016-04-27 16:38:34+00:00
2	24	00271cc10e0245fba4a35e76e669c281	San Antonio, Texas Area	Animation/cartooning	2018-09-12 16:42:17+00:00
3	4	003cc21be89d4e42bc442413a378e86	Boston, Massachusetts	Insurance	2017-10-19 18:08:19+00:00
4	4	0046ab8089c04b3a8df3f8c28621a818	Woodbridge Township, New Jersey	Insurance	2017-05-10 15:39:00+00:00

- ❖ Now showing all the plots and calculations , because title explains meaning of plot.



Summary of Professionals characteristics

People who answered but without any mail notification of question or without following tags

They are the ones who are actively supporting students and are self motivated

Lets look at 2 bar charts

1. Top Industries these professionals belong to

2. Top Locations these professionals come from

In Short we will collect all the data by using eda .

Data like

- a) How many mail's were send on posting question by learner to the professional's.
- b) How many were answered by professionals.
- c) How many professionals were present at the time of post.
- d) How many were present but didn't answered the question.
- e) How many of professional's are not active.

After getting all the Data from EDA.

We come up with our recommendation system

Here bag of words means sack like holding bag. Which will hold question words.

1). Created a function for bag of words for professionals to be able to recommend relevant questions.

2). Created a function for bag of words for questions to find similar questions asked by other students.

3). Created questions_id as parameter to my recommendation engine.

4). Finded the question date based on which we will create dictionary of questions and professional details.

5). Created bag of words of questions for finding similar questions

6). Created bag of words to find professionals whom to recommend this question

➤ We will use TF-IDF model, we will compare bag of words of question asked with professionals when question was asked.

7). created questions dictionary which existed when question was asked.

a). At first compare bag of words of questions asked and questions dictionary.
And if Data found recommend it to Learner. If not go to step 8

8). Here we will extract professional details like there profession, area of work , Date of joining, Experience .

9). And as soon as new post posted in Query block . It will compare professional details bow tfidf with questions asked in post bow.
And in result it will redirect those particular question to respective professionals.

10). Here we will set threshold value so that it will not create much misclassifying questions problem. (to avoid Un-neccesary email to professionals)

Note:- All the EDA ,calculation and proof of success will be given in ipynb file.

Notes when you build your final notebook:

1. You should not train any model either it can be a ML model or DL model or Countvectorizer or even simple StandardScalar
2. You should not read train data files
3. The function1 takes only one argument “X” (a single data points i.e 1*d feature) and the inside the function you will preprocess data point similar to the process you did while you featurize your train data
 - a. Ex: consider you are doing taxi demand prediction case study (problem definition: given a time and location predict the number of pickups that can happen)
 - b. so in your final notebook, you need to pass only those two values
 - c. def final(X):
preprocess data i.e data cleaning, filling missing values etc
compute features based on this X

```
use pre trained model  
return predicted outputs  
final([time, location])
```

- d. in the instructions, we have mentioned two functions one with original values and one without it
 - e. final([time, location]) # in this function you need to return the predictions, no need to compute the metric
 - f. final(set of [time, location] values, corresponding Y values) # when you pass the Y values, we can compute the error metric(Y, y_predict)
4. After you have preprocessed the data point you will featurize it, with the help of trained vectorizers or methods you have followed for your train data
 5. Assume this function is like you are productionizing the best model you have built, you need to measure the time for predicting and report the time. Make sure you keep the time as low as possible
 6. Check this live session: <https://www.appliedaicourse.com/lecture/11/applied-machine-learning-online-course/4148/hands-on-live-session-deploy-an-ml-model-using-apis-on-aws/5/module-5-feature-engineering-productionization-and-deployment-of-ml-models>

Link for paper Review's :-

<https://drive.google.com/drive/folders/18E9csKxRn-rz36kOFlxF59RUs3Mcg50y?usp=sharing>