



**CareerVillage is a community where students  
can get free personalized career advice from real-life professionals.**

# Creating a Question Recommender

Phase 1



# Students Can Search the Site Based on Popular Topics

[Read](#)[Ask](#)[Sign Up](#)[Log In](#) 

## Topics

### accounting

Ensuring a business is operating efficiently, an accountant does more than crunch numbers. From an internal auditor to a public accountant, there are multiple opportunities for an accountant. For more...

[View questions tagged accounting](#)

### any

Unsure of which tag applies to your question? That is fine. Just use the tags you think is closely related to your question, and apply the any to be sure.

[View questions tagged any](#)

### architecture

An architect is the first step in designing a structure. Popular careers include Architectural and engineering managers, landscape artists, civil engineers, but there is much more. For more information, please...

[View questions tagged architecture](#)

# Ask Their Own Question



Aun M. · Feb 23 · 55 views

## Career pathways for Mechatronics?

I've researched a little on the topic, but have only found videos about robotics engineers. Is this the only pathway specialized for an individual with a mechatronics degree? Or is it broader than that? As a tenth grader, I want to be fully informed of every opportunity I have in engineering...

#engineering #robotics #career #technology

↑ 4 votes

💬 3 answers

Active 17 mins ago 

📍 Ottawa, Ontario, Canada



Olivia Y. · Mar 06 · 48 views


## Where can an introverted student gain leadership experience?

#volunteer #networking #leadership #internship...

#introverts

↑ 4 votes

💬 4 answers

Active 1 hour ago 

📍 Detroit, Michigan

# Can Create a Career Path To-Do List



Renee W.

Pasadena, CA

## MY CAREER GOALS STATEMENT

To get better recommendations, [add a Career Goals Statement.](#)

Add a new To Do



## My First List (0)

+ [New List](#)

Let professionals  
suggest To Dos.

Enabled

🌐 All To Do Lists are  
public

## My First List

*To Dos that have not been assigned to a specific To Do List show up here.*

**Not sure what to put down?** Here are some examples.



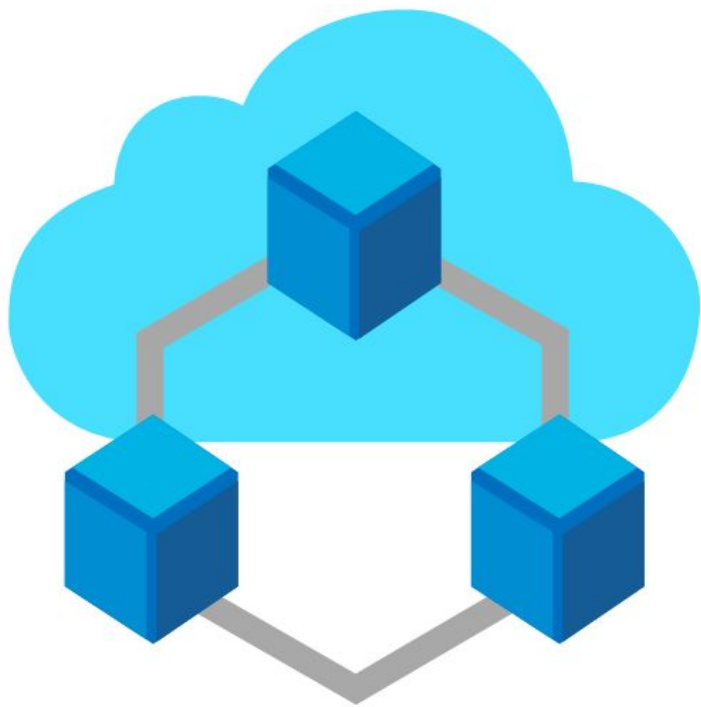
Write down five schools I would like to attend.



Discover what an astronaut's training needs to be.



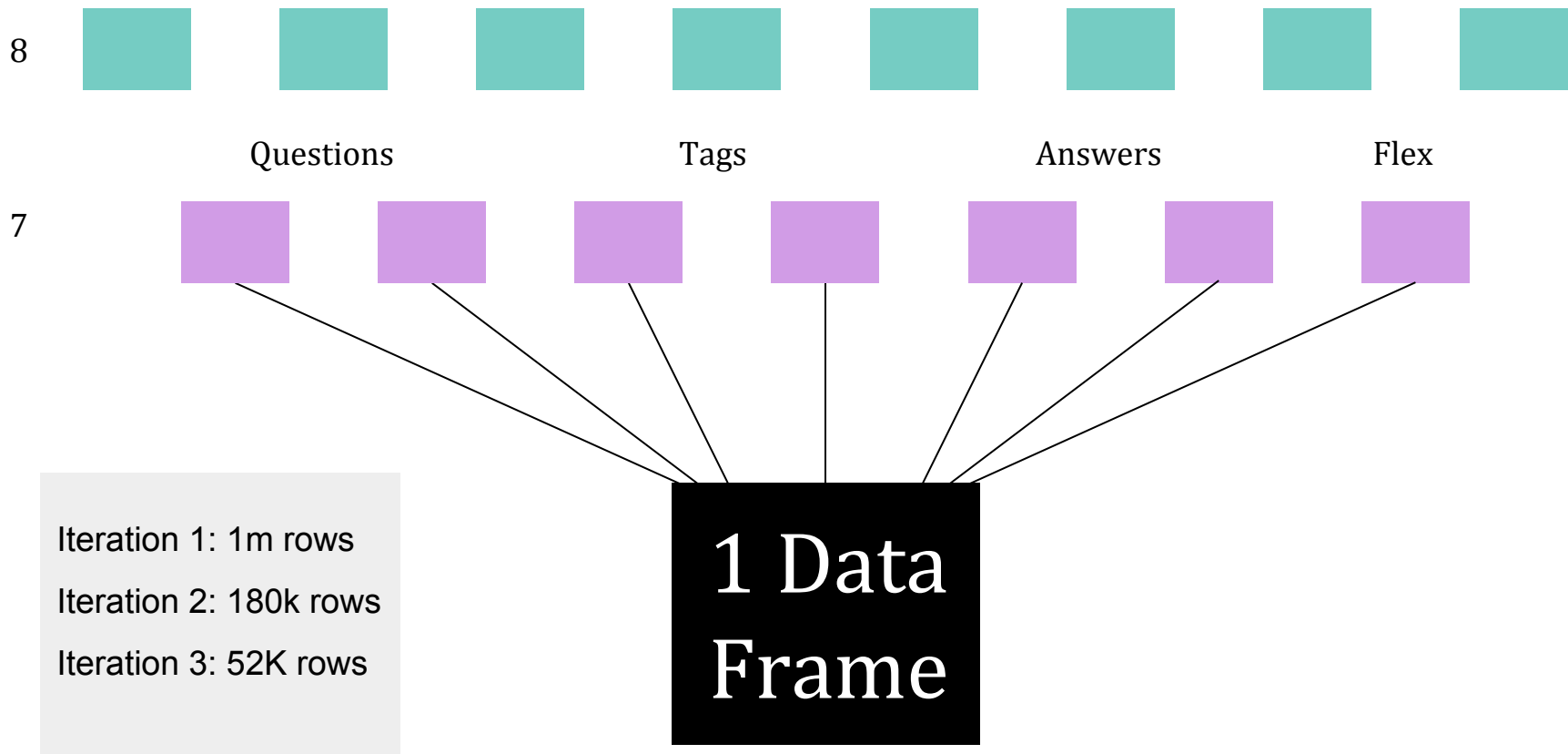
Learn the difference between software and structural engineering.



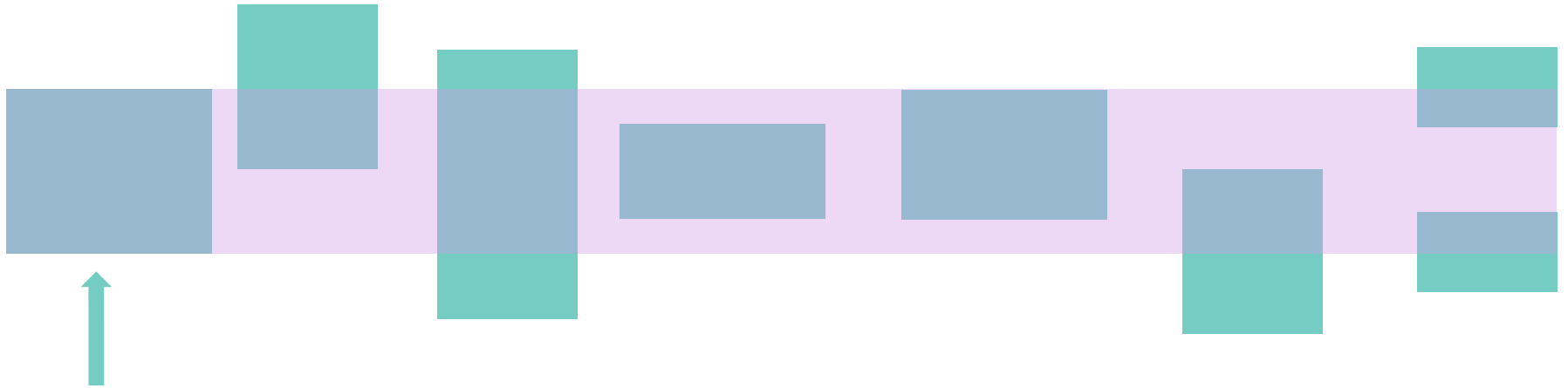
# The Data

3.5 Million Users  
25K Volunteers  
5 Years of Data

# 15 Dataframes



# Merging Madness



Questions.csv



# How Many Questions are Answered:

51123 - Answered

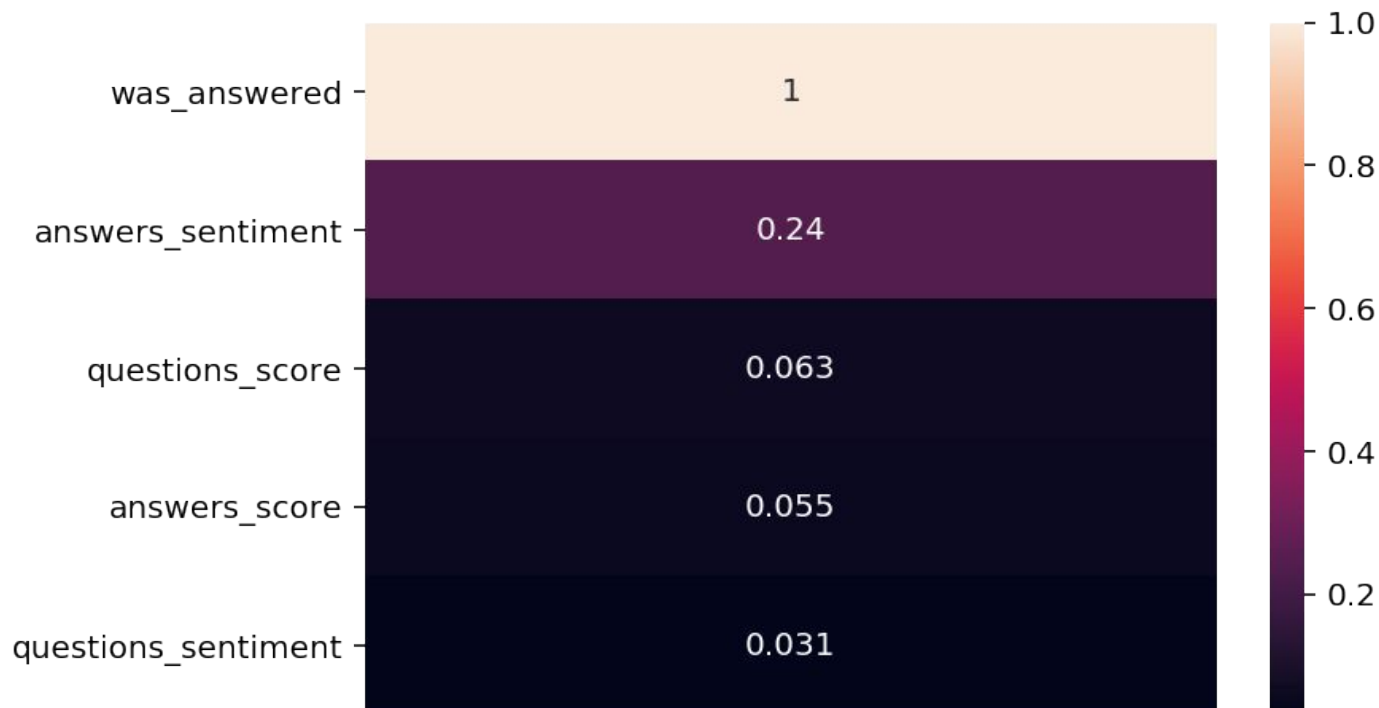
821 - TBA

51,944 - Total

98%



# Correlation to Question Being Answered

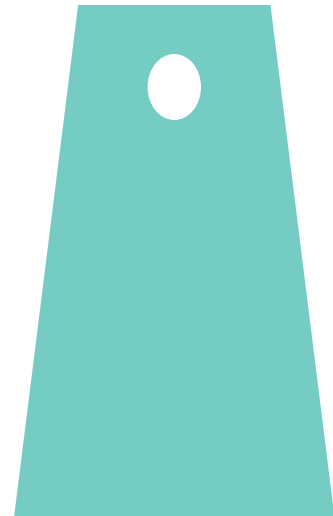




# Out of 23,288 Questions:

7091 Unique Tags

3687 Used Only Once (52%)





What does a career in data science look like?

Keep it short and simple.

I've been hearing a lot about machine learning and data science, it seems interesting. What are the job prospects like? How much money do you make? What type of problems do you work on?

#datascience #data #technology #career #datta



40 words • 5 tags

#### Suggested tags

#science

#biology

#computer-science

Use tags to match to the best professionals. (Ex: **#technology**,  
**#medicine**)

## Asking a Question Example:

- Suggests tags
- Can free-form your own tags
- Doesn't catch tag typos

Post question

# Handling Vague Questions

what should I do with my life?

Keep it short and simple.

I'm not sure what to do with my life, any advice?

12 words • 0 tags

## Suggested tags

#teacher

#future

#advice

#teaching

#nursing

Use tags to match to the best professionals. (Ex: #technology, #medicine)



# Common Tags

'college', 'career', 'medicine', 'engineering', 'business', 'doctor',  
'college-major', 'science', 'nursing', 'psychology', 'pharmacist',  
'computers', 'medical-practice', 'actor', 'coding',  
'occupational-therapy', 'biotechnology', 'fashion-design',  
'anthropology', 'film-production'

# The problem with students creating their own tag:

- One time tags

'aeronautical-engineer', 'springstart', 'futureiskey', 'fast', 'oracle',  
'interior-architectural-design', 'stl', 'health-field', 'that', 'Online-work',  
'recording-arts', 'gamewarden', 'cartoon-writer', 'cleanwater', 'onlinekicks',  
'chiropractors', 'sonographytechnician', 'orthodontics-practice', 'calpolypomona',  
'possiblelawyertoo'

# Tagging is Required

# # # # # # #



What do you need to study to become a data scientist?

Keep it short and simple.

I hear that it's heavy on stats and programming. Are there specific languages you recommend?

16 words • 0 tags

*Please include at least one hashtag*



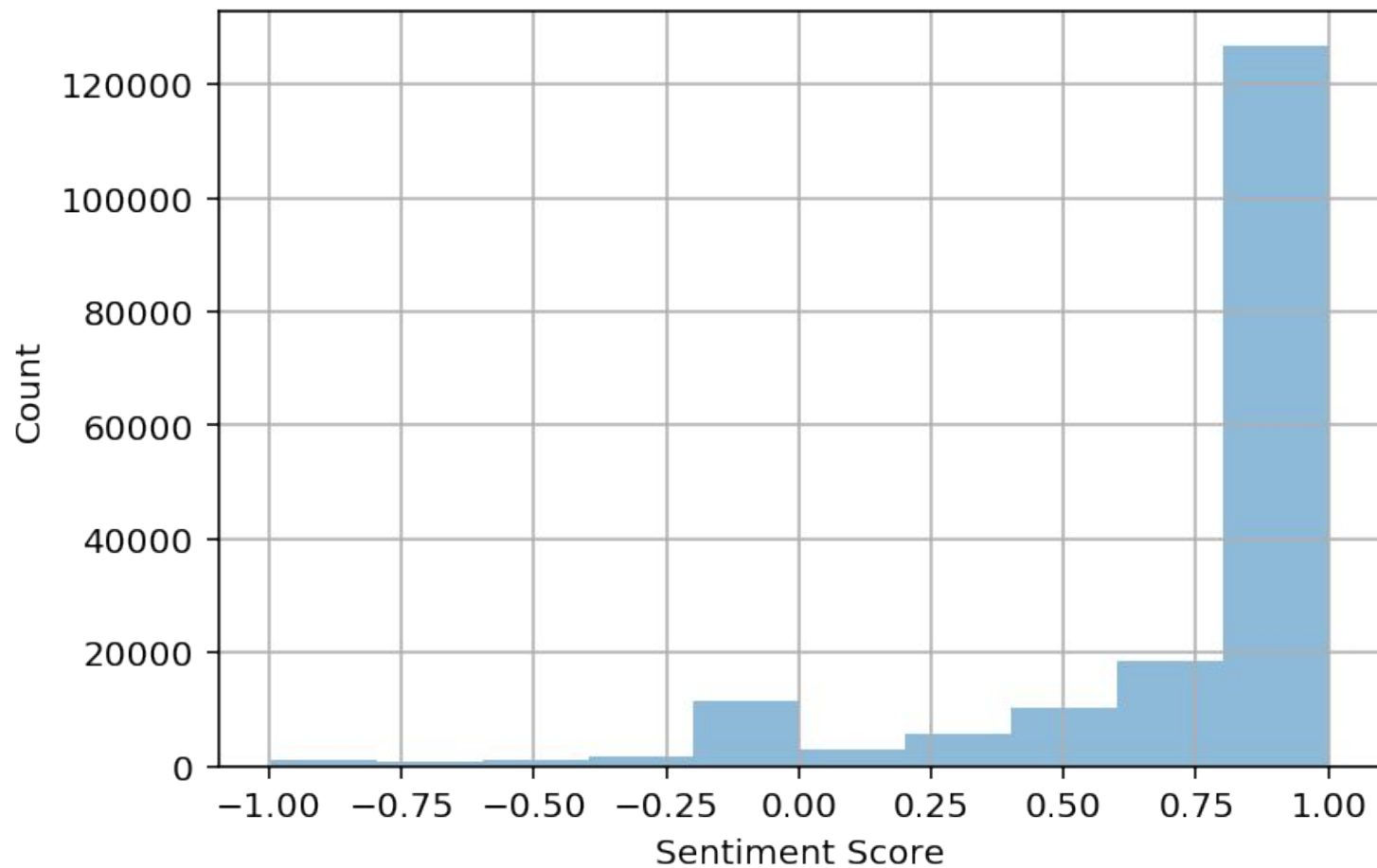
# Length of Time Till Question is Answered

Average 142 Days

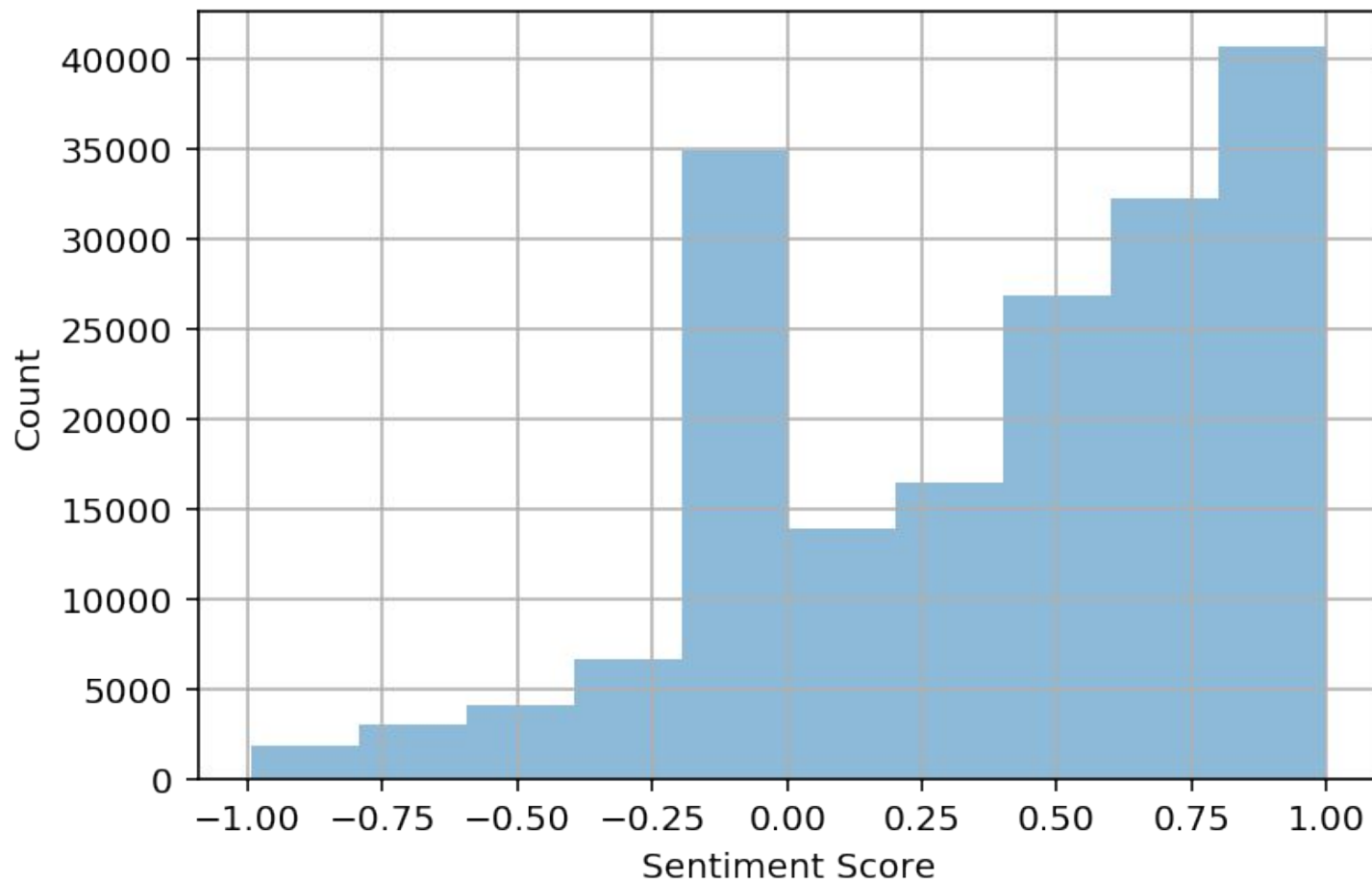
Median 22 Days

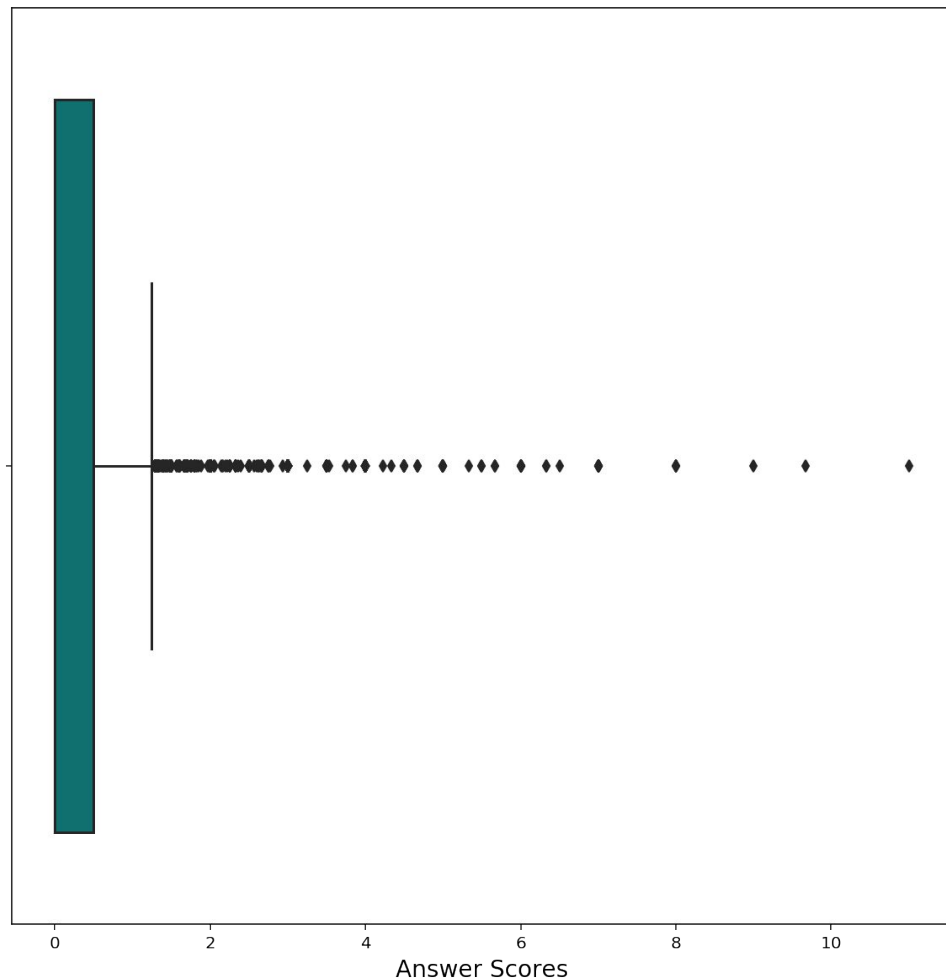
Majority 1-183 Days

Answers Sentiment Score, range -1 to 1



Questions Sentiment Score, range -1 to 1





## Professionals Avg. Answer Score

**Out of 10,169 Professionals  
Engaged:**

**1017 (10%)** above 1.00

**319 (3%)** above 2.00

**48 (0.04%)** above 5.00

Can We Predict If a Question Will Be Answered  
Based on Question Body and Question Score?



# Predicting If Question Will Be Answered:

Logistic Regression

99% Accuracy

Baseline 98%

	Predicted Wasn't	Predicted Was
Actually Was	20	171
Actually Wasn't	22	12773

# Random Forest

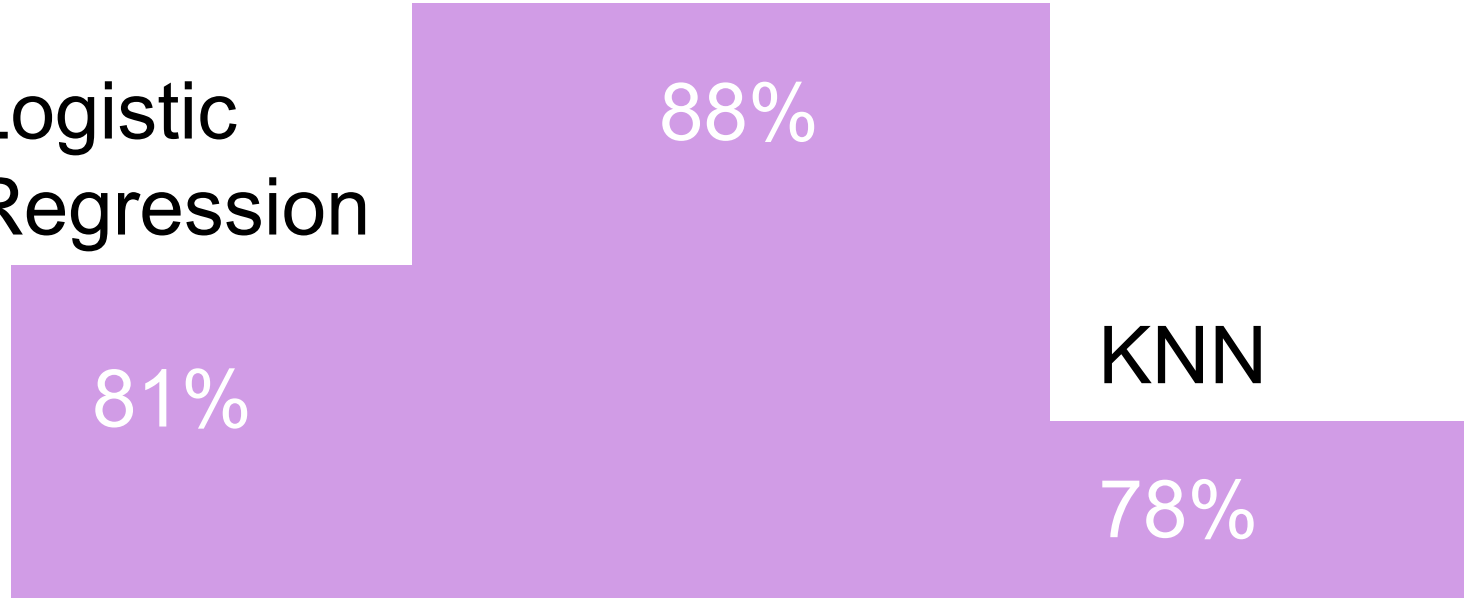
Logistic  
Regression

81%

88%

KNN

78%



## Sampled Data

Predicting If Question Will Be Answered:



Random Forest

900 Were  
821 Weren't

	Train	Test
Accuracy	100%	88%
Baseline		52%



# Random Forest Stats



Train Accuracy Score: 1.0

Test Accuracy Score: 0.88

	Predicted wasn't answered	Predicted was answered
Actual wasn't answered	178	22
Actual was answered	30	201

Can We Predict If a Question Will Be Answered  
Based on Just Question Body?



## Just Body Text

Predicting Whether Question Will Be Answered:

900 Were  
821 Weren't

### Logistic Regression

	Train	Test
Accuracy	96%	83%
Baseline	-	52%

# Features Indicative If the Question Will Be Answered Based on just the Body Text

csusmfreshman

entrepreneurship

hard

accounting

lawyer

social

military

interviews

computer

media

love

carpentry

doubt

programming

industry

career

bioengineering

assign

figures

research

**Will Be  
Answered**



**Will NOT  
Be Answered**



criminaljustice

premed

pathology

choice

minor

computers

paying

help

biochemistry

colleges

exam

veterinarian

difficult

gap

arts

physical

admissions

aerospace

tips

considering

## Phase 2

- Further Data Cleaning
- Further EDA
  - Professionals
  - Unexplored CSVs
  - Length of Time for Questions to Be Answered
- Model Time to Answered

Next  
Steps

...

# Phase 3

- Build Recommender
- Cluster Tags
- Optimize for Priorities

Next  
Steps

...



**Efficiency**  
**Vs.**  
**Community**

---

# Efficiency

---

# Community

---

Map to Qualified  
Professional



- Map Questions to Fast Responding Professional



- Map to Compassionate Professional
- Map to Local Professional



# CareerVillage Recommendations

- Data Dictionary
- Restructure How Data is Stored
- Force Suggested Tagging
- Create Tag Categories
- Forced Professional Career Suggested Tagging

Thank You



Renee Wilkening  
Data Scientist