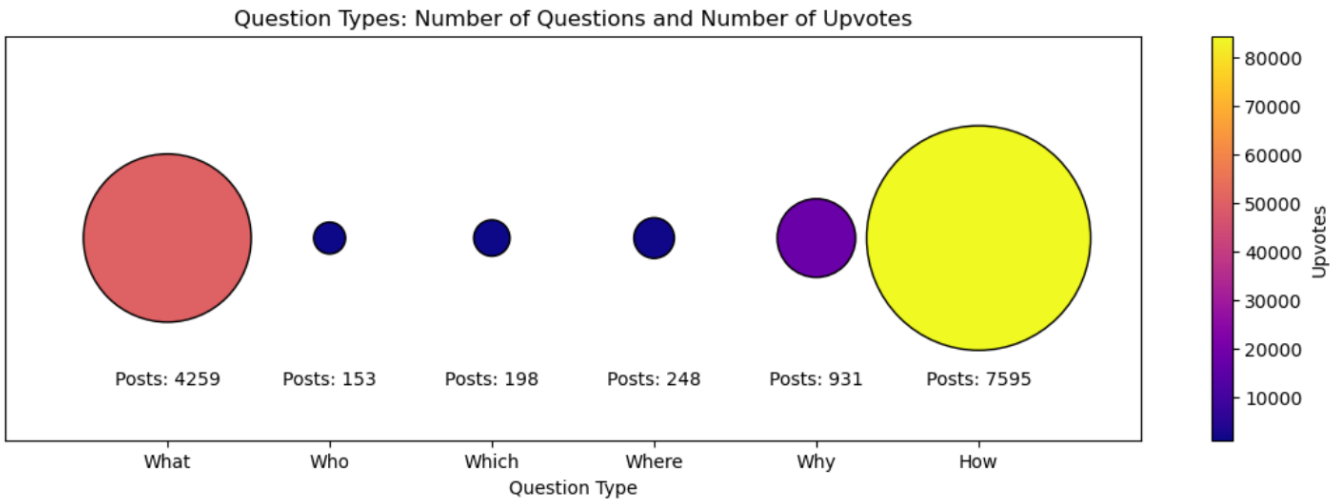
Visualisation 3:



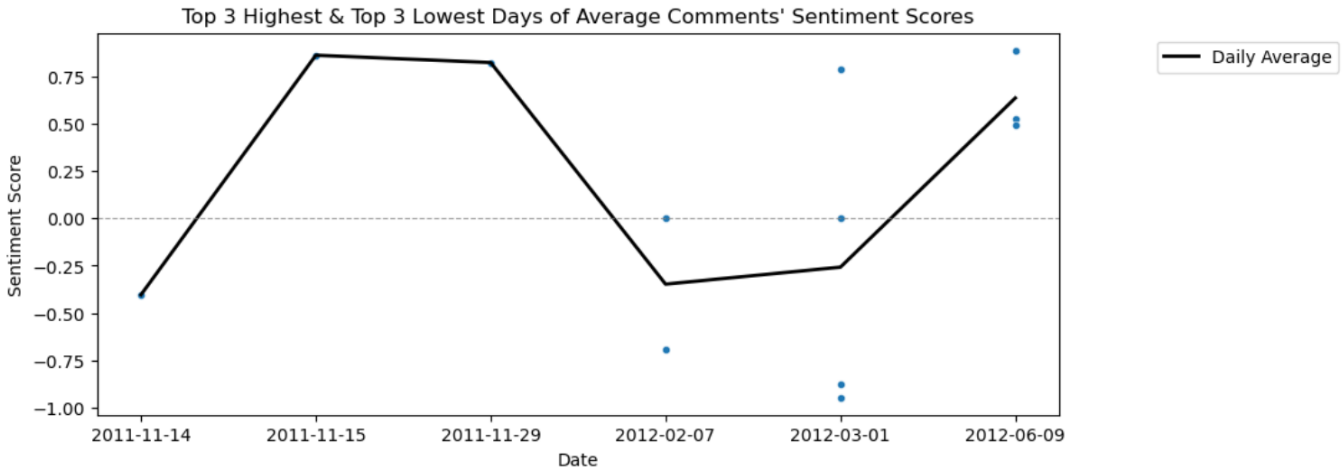

```
posts_votes_df[posts_votes_df["QuestionType"] == "How"][["Title"]]
```

	Title
15	How do I select a graduate program?
19	How many physician-scientists pursue academic tracks?
74	University rank/stature - How much does it affect one's career post-Ph.D?
82	How do you judge the quality of a journal?
98	How to find a good topic for a PhD research proposal?
...	...
134696	How do I reframe getting an MPhil instead of a PhD the first time round, in the context of applying for a mathematical postdoc?
134750	How to get a single issue of a print journal
134766	How do I motivate my PhD students to attend seminars
134778	How to properly capitalize a title in APA 7 style?
134782	How best can I explain the professional ethics of GenAI to students?

```
posts_votes_df[posts_votes_df["QuestionType"] == "Why"][["Title"]]
```

	Title
652	Why don't researchers publish failed experiments?
801	Why do universities place a weight on GRE/TOEFL scores?
1050	Why does professor prohibit me talking to people?
1084	Why do so few universities offer OpenCourseWare videos of their lessons?
1184	Why do universities have to spend money on journals?
...	...
134161	Why do most US regional colleges/universities/institutions have poor reputations in academic circles?
134220	Why has arXiv been posting articles later?
134232	Why does the status of my manuscript keep changing?
134354	Why can the same person sit on the boards of trustees of multiple schools? Has this caused problems?
134609	Why do a lot of publications ignore better results from other papers // compare only to old results (example in deep learning and biomed)

Visualisation 4:




```
(selected_data[selected_data["Date"]
               .isin(["2011-11-15", "2011-11-29", "2012-06-09"])
              ]
              [["Text", "Date", "SentimentScore"]]
)
```

	Text	Date	SentimentScore
1513	For publicly available data, transparency would dictate that code should go from the "raw data" that is publicly available to the fully transformed data. If not available via internet, it should be posted. The one "downside" to having to make data and data transformations public is that it lowers the reward to cleverly compiling/cleaning data sets if you don't have a monopoly on using them. As for confidentiality, researchers must respect confidential or proprietary data. However, in that case they should provide "example data" so that you can understand how the code worked.	2011-11-15	0.8625
1466	This answer reminds me of Richard Dawkins' description of the height of trees in a forest. The population of trees would benefit if the average height were lower but because each individual tree benefits from being taller than average, the genes for greater height spread. The size of the redwood forests does not bode well for the prospect of future grade inflation.	2011-11-29	0.8238
2304	I always tell my students to include the line "My advisor Jeff Erickson suggested that I write to you." _and_ to CC me on the email, so that the recipient sees that I can see that the student is name-dropping me, strongly implying that the student is telling the truth. (Also, I have spent the last several years building up an immunity to iocaine powder.)	2012-06-09	0.5267
2305	"Personal communication" is what you write when you've given up on citing anything that could actually help the reader, and just want to give credit to your friend. Which is about as much as citing a presentation accomplishes.	2012-06-09	0.8885
2306	Do you care more about learning stuff or making believe others that you are learning?	2012-06-09	0.4939

```
(selected_data[selected_data["Date"]
               .isin(["2011-11-14", "2012-02-07", "2012-03-01"])
              ]
              [["Text", "Date", "SentimentScore"]]
)
```

	Text	Date	SentimentScore
1512	Publishing code is one thing, but what about the data? If you are using publicly available data, then there should be no excuse for not being transparent. But what about those who are working with confidential datasets?	2011-11-14	-0.4039
2872	[Mendeley](http://www.mendeley.com/) has the groups feature, but I don't think that there is a discussion feature.	2012-02-07	0.0000
2873	You can hold a discussion in a Mendeley group, but not specifically about one paper, and the interface is lousy (think Facebook wall).	2012-02-07	-0.6956
424	In particular, do you mean personal troubles (like anxiety or depression or impostor syndrome) or professional troubles (like administrative hurdles or paper rejections or an irresponsible advisor)?	2012-03-01	-0.9468
425	See [this related question](http://cstheory.stackexchange.com/questions/3111/funny-tcs-related-papers-etc) on the theoretical computer science SE site. (Warning: lots of inside baseball.)	2012-03-01	0.0000
426	Professional troubles. But also personal troubles that arise out of professional troubles too.	2012-03-01	-0.8750
427	Given all of the suggestions to utilize text searches more effectively I might have to expand my tool-box. Thank you for the insight.	2012-03-01	0.7902