1. - When the companies are hiring
2. people for a data science team,
3. maybe a data scientist or an analyst,
4. or a chief data scientist,
5. the tendency would be to find the person
6. who has all the skills,
7. that they know the domain specific knowledge,
8. they are excellent in analyzing structured
9. and unstructured data,
10. and they are great at presenting,
11. and they've got great storytelling skills.
12. So if you'll put all this together,
13. you will realize you're looking for a unicorn,
14. and your odds of finding a unicorn are pretty rare.
15. I think what you need to do is to see,
16. given the pool of applicants you have,
17. who has the most resonance with your firm's DNA.
18. Because, see, you can teach analytic skills.
19. Anyone can learn analytic skills
20. if they would dedicate time and effort to it.
21. But what really matters is who is passionate
22. about the kind of business that you do.
23. Someone could be a great data scientist
24. in the retail environment,
25. but they may not be that exited about
26. working in IT-related firms,
27. or working with gigabytes of web logs.
28. But if someone is excited about those web logs,
29. or someone is excited about health-related data,
30. then they would be able to contribute
31. to your productivity much more so.
32. And I would say if I'm looking for someone,
33. if I have to put together a data science team,
34. I would first look for curiosity.
35. Is that person curious about things?
36. Not just for data science, but anything.
37. Are they curious about why this room
38. is painted a certain way?
39. Why the bookshelves have books and what kind of books?
40. They have to have a certain degree of curiosity
41. about everything that is in their vision
42. that they look at.
43. The second thing is, do they have a sense of humor?
44. Because you see, you have to have lighthearted about it.
45. If someone is too serious about it,
46. they probably would take it too seriously,
47. and would not be able to look at the lighter elements.
48. The third thing I think,
49. and I think the last that I would look for,
50. if I have to have a hierarchy,
51. the last thing I would look for are technical skills.
52. I would go through these social skills,
53. curiosity, sense of humor, the ability to tell a story,
54. the ability to know that there is a story there.
55. And then once all is there,
56. then I will say,
57. "Well, can you do the technical side of it?"
58. And if there is some hope
59. or some sign of some technical skills,
60. I would take them because I can train them
61. in whatever skills they need.
62. But I cannot teach curiosity.
63. **I cannot teach storytelling.**
64. I cannot certainly instill sense of humor in anyone.
65. - I think there's no hard and fast rule
66. for hiring data scientists.
67. I think it's gonna be a case by case thing.
68. I would say there has to be some sort of
69. technical component.
70. Somebody should be able to manipulate the data.
71. They should be able to communicate
72. what they find in the data.
73. I find quite often, nobody really cares about the
74. R squared or confidence interval.
75. So you have to be able to introduce those things
76. and explain something in a compelling way.
77. And they also have to find somebody who is relatable,
78. because data science, it being typically new means that
79. the person in that role has to make relationships
80. and they have to work across different departments.
81. - If this data scientist has a good mathematics
82. and statistic background.
83. - They have to consider problem solving abilities
84. and analysis.
85. A data scientist needs to be good in analyzing problems.
86. - The persons they are hiring,
87. they should love to play with data,
88. and then they know how to play with the data visualization.
89. They have analytical thinking.
90. - When a company is hiring,
91. anyone to work on a data science team,
92. they need to think about what role that person
93. is going to take.
94. Before a company begins, they need to understand
95. what they want out of their data science team.
96. And then they need to hire to begin it.
97. As they grow a data science team,
98. they need to understand whether they need
99. engineers, architects, designers to work on visualization,
100. or whether they just need more people
101. who can multiply large matrices.
102. - From a skills point, let's focus on the technical skills,
103. and in that case, first thing would be
104. what kind of technical platform would you like to adopt.
105. Let's say you wanna work in a structured data environment,
106. and let's say you wanna work in market research.
107. Then the type of skills you need are slightly different
108. than someone who would like to work
109. in big data environments.
110. If you wanna work in the traditional market research
111. structured data environment,
112. your skills should be some statistical knowledge,
113. some knowledge of basic statistical algorithms,
114. maybe some machine learning algorithms,
115. and these are the tools that you would like to develop.
116. If you wanna work in big data,
117. then there's the other aspect of it,
118. and that is to be able to store data.
119. So you start with the expertise
120. in storing large amounts of data,
121. and then you look into platforms that allow you to do that.
122. The next step would be able to manipulate
123. large amounts of data,
124. and the final step would be to apply
125. algorithms to those large sets of data.
126. So it's a three-step process,
127. but most likely it starts...
128. Most importantly, it starts with where you would like to be,
129. in what field, in what domain.
130. So, in terms of platforms,
131. let's say you wanna be in a traditional
132. predictive analytics environment,
133. and you're not working with big data,
134. then R, or Stata, or Python would be your tools.
135. If you're working mostly with unstructured data,
136. then Python is more suitable than R.
137. If you're working with big data, then Hadoop and Spark
138. are the environments that you will be working with.
139. So it all depends upon where you would like to be,
140. and what kind of work excites you,
141. and then you pick your tools.
142. In additional to technical skills,
143. the second aspect of the data science
144. is to have the ability communicate,
145. the communication skills or presentation skills.
146. I call them storytelling skills.
147. That is that you have your analysis done,
148. now can you tell a great story from it?
149. If you have a very large table,
150. can you synthesize this and make it more appealing
151. that when it goes on the screen,
152. or is it part of document that it just speaks,
153. it sings the findings,
154. and the reader just gets it right there?
155. So, the ability to present your findings,
156. either verbally, or in a presentation,
157. or in a document,
158. so that communication and presentation skills
159. are equally important as the technical skills are.
160. When you have a great insight,
161. and when you're presenting your results...
162. Imagine you're driving on a mountain
163. and then there's a sharp turn,
164. and you can't see what's beyond the turn,
165. and then you make that turn,
166. and then suddenly you see
167. a tremendous valley in front of you,
168. and this great sense of awe that,
169. "I didn't know that."
170. Right?
171. So when you present your findings
172. and you have this great finding,
173. and you communicate it well,
174. this is what people feel because they were not expecting it.
175. They were not aware of it,
176. and then this great sense of happiness that,
177. "Now I know. And I didn't notice. Now I know."
178. And then it empowers them.
179. It gives them ideas what they can do with this knowledge,
180. this new insight.
181. It's a great sense of joy.
182. As a data scientist, you're able to share it
183. with your clients because you enabled it.