Data Dames: When does a job post go stale?

We were interested in using a senior college student's insight to shape how employers could improve upon job listings. These variables were chosen based on an applicant's consideration of industry type, job location, job qualifications, position description, and age of job post. We found that this last variable of how long the job had been posted was a crucial variable for employers to consider, especially if they are paying Indeed to be sponsored.

Indeed differentiates "organic listings" from "sponsored jobs" by making the sponsored jobs remain high in the search results and appear prominently at the top and bottom of the page even as time passes while organic listings fall in the search results due to new jobs that are added at a rate of 8.2 per second (blog.indeed.com). While we did not have the data to determine which companies were sponsored, we were able to use this knowledge for later interpretations.

We tried to predict the number of applications with simple linear regression, multiple regression, and mixed models. We wanted to find which were the factors that lead a job post to having more application submissions in Indeed. After analyzing our regression models and seeing very low r-squared values, and a high sensitivity in the p-values, due to the size of the data, we went back to analyzing the data from a visualization point of view. We plotted multiple variables against the number of application and we noticed that a common denominator was time. Around the 4 week mark, no matter what the variable was, the number of applies for job posts start go stale. This is what lead up to looking at the number of applications per week, over the first six weeks a job posting has been on Indeed.

Taking a random sample of 10,000 jobs, we created six different boxplots, one for each week after a job posting has been created, to compare how the number of applications changed over time. These plots showed that a bulk of the number of applications happen in the first and third week age of a job posting. It also showed that after week three the number of applications per week starts to show no change and the biggest difference was the number of outliers for each box plot. Basically, the same companies consistently receive extremely high number of applications, within the first 6 weeks. However, as the number of weeks increase, the number of outliers decrease. This might explain why the total number of applications goes stale around the 4 week mark. One reason for these outliers might be due to certain companies having their job postings sponsored. Sponsored Jobs are based on a pay for performance model. Employers set a monthly budget and only pay when a candidate clicks to view your job. You decide how much to spend and we deliver relevant, high-quality candidate traffic. Some other confounding variables might include: new users on Indeed, company fame, and type of industry