INTRO---INDEX PAGE

This project grew out of a major pivot point in my own life. I spent most of my career as a human resources manager and consultant, but last summer I decided that what I really wanted to do was focus on the aspect of my practice that interests me the most, and that is data: analyzing data to help management figure out solutions that are based in hard fact, not just instinct or so-called best practices.

So in my data science studies I’m always thinking about how I can apply what I’ve learned to people management situations. This project is an attempt to do that. I downloaded a Kaggle dataset with information about nearly 1500 current and former employees in a fictional company. The goal was to build a dashboard that would help managers understand the drivers behind employee turnover and where they should be focusing their efforts to curtail it.

METHODOLOGY

I wanted to mimic as much as possible how the app would work in the real world, where data would be continually changing, so I loaded the Kaggle csv into a PostgreSQL database. All the calculations in the app are being drawn from the database so they can change in real time.

The base data is cleaned and calculated in python and served via flask to the web app. The tools in the app are built with javascript.

So on the home page you see a broad overview of the data:

Overall turnover rate and then the results of a logistic regression to determine which features in the data have the most impact on the outcomes.

DATA TOOLS

The user can then dig successively deeper into the data. [**data tools**] This first tool allows the user to look at individual features of thedata--what I’ve called here “employment factors”--and see how they are correlated with turnover. For example, if we choose age, we see there is a general decrease in turnover (the orange bars) as age increases. However the table gives us a little more background data so we can see that there are only 5 individuals in the over 60 category so that data is probably less reliable. But the fact that employees under 30 are more likely to leave by over 13 percentage points is something that bears looking into.

COST

The next tool allows managers to explore how important this turnover might be to the bottom line. Current turnover rate is 16.1 percent and the estgimated annual cost of that turnover is $1.4 million. The user can put in a hypothetical rate and see how much the company would save—or how much more it would lose. For example, if we reduced turnover by, say, 5 percentage points to **11.1** percent per year, we’d save close to half a million dollars.

The prediction tool allows managers to look at individual employees they may be concerned about. So you enter an employee number and the tool gives you the odds that he or she might leave the company. [ENTER **776**, THEN **53**.] If you enter a number not currently assigned to an active employee [**365**], you get a notice to that effect.

DEEP DIVE

At the next level of analysis, users can explore a number of different visualizations. These are curated by the analyst—that is, me—based on relationships that seemed most meaningful and rendered in tableau. (Theoretically, this data would also be drawn from the database and updated in real time, but the public version of tableau doesn’t allow that functionality, so these charts are drawing data from the csv file.)

So for example we can look at a simple chart on **gender**: Compared to their representation in the workforce (the chart on the left), women are less likely to leave than men. Earlier we saw that younger employees left at higher rates. We might wonder how those two factors interact. [**age and gender]**

Men are represented in these charts by the darker colors. So it turns out that while women have lower turnover rates overall, at certain points—between 20 and 25 and between 45 to 50, for example, they make up more than half of attrition.

In the set of charts below, we can also explore the data by department. [**turnover by department**] (This dataset only presents three departments—R&D, sales, and HR). It turns out HR and R&D folks leave at considerably higher rates than their share of the workforce. So let’s look at how gender, for example, play into that: [gender] Men are orange—in sales , men are leaving at considerably higher rates than women but in HR women are much more likely to leave than men---something that warrants further investigation.

RESOURCES

The last page of the app points managers to some resources that might help them think through strategic responses to the data: They can explore the company’s data to regional and industry=wide experience, and they can view media and articles about how to encourage employees to stay.