

DATAFEST 2018

—

KUANJU CHIU, XIAOKE MENG, JASON LI, ZELIN LI

Dataset: Indeed Job Listings

Variables used	
avgOverallCompanyRating	salaryCurrency
numOfRatings	supervisingJob
normTitle	licenseRequiredJob
normTitleCategory	educationRequirement
descriptionWordCount	Average clicks per day
experienceRequired	
estimatedSalary	

I. HOW TO MOVE BETWEEN INCOME BRACKETS?

- WE ESTIMATE THE FACTORS THAT COULD AFFECT EACH INCOME BRACKET
- Convert to USD, remove estimated salaries = 0 (2%), education -> binary, corrected values
- assumptions: full time jobs (>15000), {med, admin, management} => 1
- Partition salaries into three parts: 15 000 – 50 000 LOW, 50 001 – 100 000 MID, > 100 001 HIGH
- Feature selection with Pearson method

■ LOW:	License Required	Experience Required	Supervising Job	Number of Reviews
	0.170478	0.116662	0.10424	0.081644
■ MID:	Experience Required	License Required	Education	Supervising Job
	0.283208	0.178312	0.128825	0.123302
■ HIGH:	Education	License Required	Experience Required	Supervising Job
	0.272652	0.262715	0.135605	0.113997

