

Visual Analytics System for Job Analysis

Sai Nikhil Gundu, Ramesh Rajan, Megha Devrani

Abstract: The goal of the analysis is to provide jobseekers with job-related information which can guide them when searching for specific job openings. Especially, those who long for a change in their career may want to have clear idea of candidate location or positions to consider. The visual analytics system is intended to support the exploration and decision making for new jobs. The system is built based on the dataset of more than 30,000 federal government jobs, using D3 along with HTML, CSS, dc.js and Javascript, to develop visualization and interaction.

Index Terms— Job Search, Visual Analytics, D3 Implementation, dc.js

Introduction

The internet is a prominent source for American job seekers. It was in 1990s that job search engines came into effect ^[1]. Most of job searches were based on the internet searches ^[2]. The system is designed to allow the job seekers to go online to check for the job information. These job websites allow employers to publish about their jobs and match the relevant job seekers. They range from just job descriptions to some more relevant information like pay statistics and location information in plain text.

Although the job search websites are good to go resources for the job seekers with list of companies and available jobs listed, it is high time to use some intelligence system to give more insights into the job requirements. Thus, aiding the job seekers in more relevant way than the traditional descriptions. In this paper, we outline a visual analytics system that provides the capability to the job seekers about exploring more into the job market and find relevant states that they can apply for. This system also helps the job seekers with relevant information about top occupations in the department of interest and the pay information across the U.S. states.

1. Related work

Building a system about jobs entails understanding of online job seekers and their needs. Research shows that most of the people who use the Internet are already employed and are looking for a job change. It also reveals that many of them are likely to switch from current job to a new one¹. It provides

theoretical support for developing an employment website which allows the user to explore the new jobs that they do not know. Also, the job hunters will be able to compare the newly found jobs to their current one with the knowledge obtained from the system we develop.

Employment websites vary in information they deal with. Most of the websites in research like glassdoor.com, monster.com, Indeed have the information about jobs based on the location input that user gives. It is highly appreciable if these websites are more insight driven.

These websites that provide a high-level information about the occupation and overall job market. Some websites like datausa.io and Bureau of Labor statistics have the data about the occupation and department across the states. This is the dataset that can be leveraged to provide the insights into the system and maintain more transparency. Our system takes advantage of this raw data and power of visualizations to build such robust job analysis system.

2. System design

The design goal of our job analysis system is to find the jobs within the department of choice across the U.S. showing the heatmap of job concentrations, different occupations in the department, top occupations and top paying jobs with state information.

The dashboard is a single page with 6 visualizations speaking about different scenarios that user can interpret. It has the dataset read from d3 with quick results upon filtering. The corresponding department or occupation can be

clicked graphically to interact with other visualizations present in the dashboard. There is a reset functionality provided for each result to go back and check the information in case user needs to. The data is sorted with respect to mean annual salary in tabular forms to give more information about highest and lowest paid occupations with states.

2.1 Database Selection

The two datasets are used for the system: the dataset of 'Annual and Weekly wage of all government employees for the year 2016' from the United States Bureau of Labor Statistics and the one of 'Household Income: 2016' from United States Census Bureau. The former one contains information about 34,401 government jobs consisting of the profile of job, the number of jobs in the given state, and the weekly and hourly wage of the job as a mean, median, and range of percentile. The latter dataset contains the median household income and Gini index in 2015 and 2016 by state. To enable the users to filter a job or a group of jobs, only relevant attributes are chosen from two datasets.

2.1.1 Database Selection and Normalization

To categorize the jobs, the occupation code and occupation title from the first dataset were selected. Since the names of department and specific position are placed in one column, occupation code was used to distinguish them.

For geographic classification of the jobs, state abbreviation and state name in the first dataset were used. The state names were matched to the code information provided and a new column with state abbreviation is created to match the state names. The mean annual and median annual salary columns are converted into numeric for use.

2.1.2 User Flow

The job analysis system we built consists of a single page used as dashboard. The user by default can see all the jobs in the dataset by default. A legend is provided beside the first pie chart to click the interested department. Once the user clicks the relevant department a heat map shows the concentration of the department jobs

with more blue indicating more number of jobs. User can thus select the corresponding state that he is interested in or with having the mouse over the state, the number of jobs present in the state is displayed.

The user also sees a pie chart that shows all the occupational titles within the department with the area representing the number of jobs with the occupational title. The user can select a particular occupation to see the statistics like top paying and low paying state information with annual mean and median salary.

With the user selection of the department, the user is enabled to see the top 15 occupations sorted by number of jobs as a histogram. This enables user to think about the occupation that he can target.

By clicking an occupation, the heatmap changes to the concentration of jobs with the occupation title. The user can select the occupation via pie chart or from the histogram provided.

The tabular displays are for the informational purpose regarding the states and salaries. User has a provision to hover on the particular row to highlight the job of interest.

Overall, the system is robust in taking the inputs on any visualization thus changing the user flow accordingly.

2.2 Visualizations

The job analysis system default page is shown in Fig 2.2.1-1.

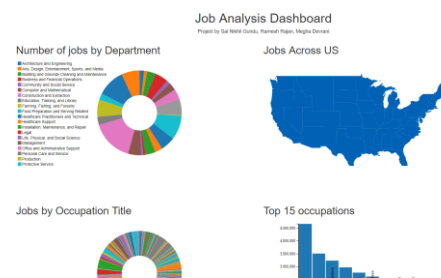


Figure 2.2.1-1. Job Analysis System

The page consists of six visualizations: two pie charts with department and occupational titles, one histogram with top occupations in the department, one heatmap showing states with

This scenario is helpful to the user in selecting the states that he can look for the department (his choice) jobs as shown in fig.2.2.1-2.

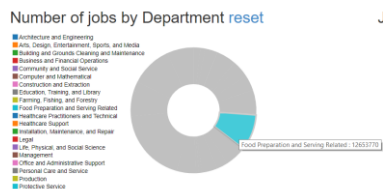


Figure 2.2.1-2. Selecting food department jobs

The different occupations within the food department with top 15 occupations with respect to the number of jobs is shown in fig.2.2.1-3.

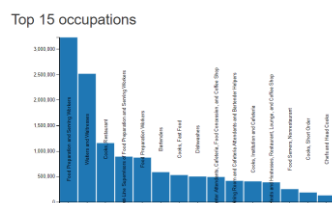
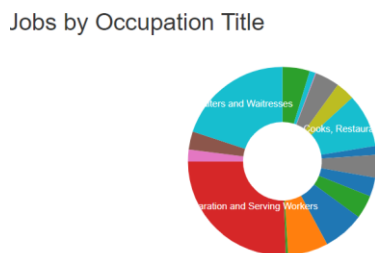


Figure 2.2.1-3. Different occupations in the food department

The user selection of food department also gives tow tables with highest and lowest paid jobs as tables. The tables show the information about the diffent occupations with states and the mean and median annual salary for the occupation. This is as shown in fig 2.2.1-4.

Top Paying occupations

STATE	OCC_TITLE	A_MEAN	A_MEDIAN
District of Columbia	Chefs and Head Cooks	60,820	53,210
New Jersey	Chefs and Head Cooks	58,200	56,730
Massachusetts	Chefs and Head Cooks	55,840	52,580
Florida	Chefs and Head Cooks	54,820	50,260
Rhode Island	Chefs and Head Cooks	53,480	52,950
Washington	Chefs and Head Cooks	52,700	47,210
Maryland	Chefs and Head Cooks	52,140	46,890
Nevada	Chefs and Head Cooks	51,380	47,500
Nebraska	Chefs and Head Cooks	51,340	44,520
West Virginia	Chefs and Head Cooks	49,100	45,370

Low Paying occupations

STATE	OCC_TITLE	A_MEAN_A_MEDIAN
Puerto Rico	Cooks, Short Order	16,910 17,450
Puerto Rico	Food Servers, Nonrestaurant	17,070 17,630
Puerto Rico	Cooks, Fast Food	17,100 17,640
Puerto Rico	Food Preparation and Serving Workers	17,110 17,540
Puerto Rico	Hosts and Hostesses, Restaurant, Lounge, and Coffee Shop	17,150 17,570
Puerto Rico	Dishwashers	17,270 17,630
Puerto Rico	Waiters and Waitresses	17,310 17,590
Puerto Rico	Food Preparation Workers	17,330 17,590
Puerto Rico	Counter Attendants, Cafeteria, Food Concession, and Coffee Shop	17,340 17,700
Tennessee	Cooks, Fast Food	14,770 17,750

Figure 2.2.1-4. Highest and lowest paid jobs in food department

2.2.2. Geographic Distribution

Under the Geographic Distribution tab, the user is presented with a choropleth by state name. This bounds every state in a polygon, and each state region is colored according to the job counts for that region. Darker regions indicate a higher number of jobs. This allows the user to see if there are any distributional trends by geography. For example, if the user selects the food department the heatmap is as shown below in fig. 2.2.2-1.

Jobs Across US

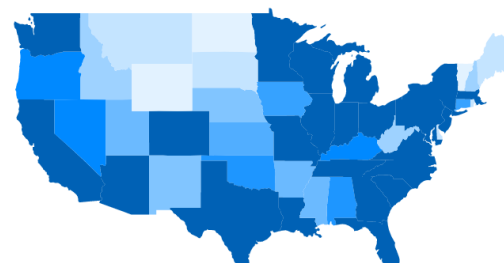


Figure 2.2.2-1. Heatmap showing job concentration

It can be seen that there are more number of food department jobs in CA, TX, etc. compared to OR.

3. Evaluation

As an implementation, the job analysis system allows the users to explore the state and job title that are best matched with their criteria in terms of kind of job, salary, and job demand. The system assists the users with various visualization to narrow down the jobs step by step. However, there are many caveats which concern the limitation of data and unrealized functions and features in the system.

3.1.1 Data Enrichment

The dataset that the system is based on contains the federal government jobs only. Therefore, the users who look for jobs in the private sector may not be able to discover the jobs of interest or have a chance to obtain incorrect information due to the difference between citizen sector and public sector. The same issue could happen to those who seek for the jobs within city government. To serve users with diverse needs as described, the data including the jobs from private sector and city government are needed.

In addition, the information about expected cost of living could be more precise and specific. The median household income of each state was chosen for that purpose since the jobs in the dataset were grouped by state. However, the price of factors deciding the cost of living is different from city to city in one state so is hard to generalize. Combined with the job's location attribute to city-level, the city-wise living cost would better inform the users. Also, wider-ranging information such as property price would enable more realistic estimation on how much the user would spend if he/she moves to the city or state offering the job found on the job analysis system.

3.1.2 Implementation that was not achieved

On the system, the users cannot see the legitimate legends for narrowing down. It is also lacked with the state names being displayed on the heatmap.

Furthermore, due to the time limitation, the filters for the search as a dropdown option is not implemented. For future improvement, the system should allow the users to choose department and job titles in the search result

page and filter out department in the search result page of keyword plus state.

In terms of functionality, improvement can be done with respect to zoom in the heatmap for a county wise job concentration.

4. Conclusion

The job analysis system is built to let the user browse through the existing jobs and find the right job for him/herself with visualizations and interactions. Since the system provides information not only in a single page but while the user is comparing and considering many jobs, the process of finalizing the most suitable job itself is illuminating. For example, a computer systems analyst living in South Carolina can learn how much he/she would learn and spend if she moves to California with her current job or how hard it would be to switch to a position of web developer in South Carolina based on the number of existing jobs from this system. Also, the user may get to know the new jobs that he/she didn't know before.

However, the system has room for improvement regarding input dataset and implementation of functions and features. Adding the data from private and more public jobs and advancing the filters and visualizations are expected to be applied for the users with diversified needs and better usability.

REFERENCES

- [1] Beard, T. Randolph, George S. Ford, and Richard Saba. "Internet Use and Job Search." SSRN Electronic Journal (n.d.): n. pag. Web.
- [2] Peter Behr, David Segal, "High-Tech Turks Lure Big-Buck Backers; Outside Investment Grows but Trails Other U.S. Centers", The Washington Post, Nov. 4, 1996, pg. F05
- [3] Pew Research Center, November, 2015, "Searching for Work in the Digital Era"