OPIM-5501-Visual Analytics-SECMS40-1228

Project Report: Analytics Job Guide - 2022

Prepared for: Analytics Majoring Graduates.

Submitted by – Team 6

TABLE OF CONTENTS

.

1. Introduction

Every year millions of students complete their educational journey and enter the work force eager to make an impact in their field. One of the greatest challenges of this is finding the right place to begin or continue this journey. With the seemingly unlimited amount of position, company and location options, graduates can become overwhelmed trying to make the right choice. Students must consider the salary, career progression, corporate values, and price of living, among other things, for each job they apply to.

On top of the coursework, they must complete to graduate in good standing. Hence, many students do not have the time or energy to properly research the job market. This project's goal is to streamline that process, so students will be able to –

- Make an informed decision on where to continue their career journey
- Find the details around the compensations offered across different roles
- Understand the career trajectory for different data-specific roles
- Understand what to look for in terms of corporate values and diversity and equity
- Learn which locations have the highest average salaries, understand the huge impact of cost of living

Using this information, students can apply, interview, and accept offers with the peace of mind knowing they made the correct, informed choice. Having this peace of mind will allow students to focus on completing their studies without having to deal with as much of the added stress and uncertainty that comes with making a big career choice.

2. Datasets

We sourced our data from **Levels.fyi** through the Kaggle repository. We acquired the data relevant to the jobs available in the Analytics field for the past three years and performed the exploratory data analysis (EDA). According to our requirements –

- We created our visualizations to show the current analytics job market
- Repository of IN-DEMAND skillsets preferred by employers
- Job Openings across the different roles and industry types
- Career Progression – Examples of specific roles
- Salary and cost of living analysis across the United States
- Industry-wise top firm compensation across generic data roles
- Demographics in the Analytics profession

For our data we have set perspectives from both employers and employees' eyes, what are they looking for, the key factors that will help employees to stay in the company for long term. Job satisfaction, employee friendly environment and such similar areas of interest.

We used datasets (see references) in across different dashboards to convey interlinked information - For example we are taking into account the cost of living in major cities across the US and determining the best place to live with respect to the compensation being given. This information can help a candidate choose a location with respect to their compensation.

MAIN DATASET:

Data Science & STEM Salaries - 2020-2021

- 62,000 salary records from top companies were pulled across every data and STEM related roles i.e., Data Analysts, Data Scientist, Management Consultant, Product Manager, Data Architect, etc.
- The parent data was scraped off **LEVELS.FYI** – which is a repository of validated salaries posted by employees
- The data included the job profile, company, location, details of compensation (Total, Base), Experience Level of the employee, Remote Ratio for the firm, Year of Data

3. Methodology

In this project we sought to find sufficient data on the analytics job market, so that we can discover which areas, geographically, career and corporate, would be best suited for graduating analytic focused students. We gathered data from various sites, such as Kaggle, government sites, and job sites. After finalizing our first few datasets, we analyzed them to make sure they were sufficient, and relevant enough to use for our project. This included making sure the data was recent, that there was enough data, and that the data as relevant and useful for our target audience. As our story developed, we kept appending onto our initial data repository to have a holistic view of the **Analytics Job Market**.

After finding datasets we could use for our project, we cleaned them using python and performed the basic Exploratory Data Analysis which included finding missing values, imputing, or removing them, seeing if data is bias towards one side. Then, we evaluated whether outliers were relevant or simply a statistical oddity and had to transform data such as currency so that they are all reflected in the same currency (USD). After cleaning our initial dataset, we followed similar processes with supporting datasets we used in the project.

Once the data was ready for tableau, we focused on building visuals based on two principles

- Useful information for the target audience
- Aesthetically pleasing dashboards that keeps them engaged

It was important to not overwhelm the audience with too much data, but also not to bore the audience with redundant or repetitive information. After building a diverse set of visuals to convey our information to the audience, we grouped them together into dashboards based on categories such as cost of living, career progression, diversity and equity, and salary, among others.

Lastly, we inserted the dashboards in a logical order onto our storyboard. To do this we made sure to put dashboards that complimented each other close to each other on the storyboard.

We also inserted brain breaks, not only to avoid a data dump in front of the audience, but also to effectively transition to a new topic in our presentation.

4. References

1. *Data Science and STEM salaries - Levels.fyi, 2021* [1]
2. *Forbes - Louis Columbus - Analytics Demand - 2020* [2]
3. *Top 5 Data Analyst skills that employers want in 2022, Stephan Greet* [3]
4. *How to become a Junior Analyst/Data Scientist - Glassdoor* [4]
5. *Cost of living in the United States - 2022* [5]
6. *Business/Data Analyst Demographics in the US - 2022* [6]
7. *Housing Dataset - Zillow* [7]
8. *Data Scientist Salary Map 2021, Galvanize* [8]
9. *Data Science Skills Survey, Kashyap Raibagi - 2022* [9]
10. *Hays US Salary* guide - 2022 [10]