Information Visualization Project: Analytics on Career in STEM

Ankita Aggarwal, Priya Mishra, Aditi Agrawal

December 2, 2017

Abstract

The project aims to guide students and working professionals who are looking for direction towards choosing a career path, more interactively. Higher focus is laid on STEM viz. Science, Technology, Engineering and Math, and many dimensions such as future career, occupation growth, employment and wages etc. were considered for the analysis. The project emphasizes on good development practices that help its readers grasp the crux of each dimension easily, by looking at visualizations on each page. The website provides an overview of STEM work, analyzing nearly 100 occupations, including data showing selected STEM occupations with the most employment and projected job openings and growth.

1 Introduction

STEM is a curriculum based on the idea of educating students in four specific disciplines — science, technology, engineering and mathematics — in an interdisciplinary and applied approach. STEM offers more work prospects for students, a better quality of education, and a capable and competent skilled workforce for the future.[ste14] As our world becomes more focused on science and technology, it is important to be prepared to keep up with the demands while ensuring your employment opportunities.

According to the U. S. Department of Commerce, STEM occupations are growing at 17%, while other occupations are growing at 9.8%. STEM degree holders have a higher income even in non-STEM careers. Science, technology, engineering and mathematics workers play a key role in the sustained growth and stability of the U.S. economy, and are a critical component to helping the U.S. win the future. STEM education creates critical thinkers, increases science literacy, and enables the next generation of innovators. Innovation leads to new products and processes that sustain our economy. This innovation and science literacy depends on a solid knowledge base in the STEM areas. Most jobs of the future will require a basic understanding of math and science. Despite these compelling facts, mathematics and science scores on average among U.S. students are lagging other developing countries.

1.1 Problem Statement

The problem we are trying to address is the fact that young millennials and working professionals are often confused as to what career path they should chose out of all the available options. For someone who is inclined to take up a career in STEM, or someone who wishes to know more about various fields in STEM and future scope, we tend to provide a single platform where we explore various types of information on STEM occupations displayed in the form of visual graphics and charts making it more intelligible and at the same time, more interactive for the user.

This way the user gets to have an overview of the opportunities that STEM offers, in more depth and since the data is more graphical than textual, it helps its user retain information for a longer time as compared to chunk of texts which take time to comprehend and may lead to user losing interest. All currently existing information sources for STEM have either textual and lengthy descriptions of the same information or non-interactive data in charts.

1.2 Data Source

Data for accomplishing the project was taken up from the Occupational Employment Statistics (OES)1, a program producing employment and wage estimates annually for over 800 occupations.

These estimates are available for the entire nation as well as for individual States, metropolitan and non-metropolitan areas; national occupational estimates for specific industries are also available.

1 A subject covered under Bureau of Labor Statistics, by United States Department of Labor, https://www.bls.gov/oes

1.3 Previous Work

The website here, provides a static analysis of the STEM careers with minimal interactivity. The data used for analysis was last updated in 2013, and does not present latest figures.

The pwc website lists the requirement for STEM graduates and the career path as a full speed forward pathway. It provides more of a marketing and hiring perspective, as perceived at pwc.

The webpage here provides several different analyses using STEM Education Data with 2012 metrics. Some points mentioned are who should pursue a STEM career, what courses to have taken in middle school and high school, education level of parents, 12th grades, number of people enrolling in different undergraduate courses, retention rates of students etc. While the scope of analysis builds the groundwork, but it might end up confusing its reader more, as to, say, if a person has not taken certain courses in his high school, he might just drop the idea of pursuing a STEM career. Looking at the attrition rate of students in undergraduate STEM courses, the reader might consider this as a horrific option with a lot of complexity.

Our website aims to overcome all the drawbacks and provide higher interactivity to user with the latest data.

2 System Description

The layout of the website is set with a navigation menu for easy movement across pages. The content on each page is set as- chart constructed from data, and the linked description and analysis of the chart alongside. The website provides our suggestions and lists the data source and the team information in the footer section.

The various sections created in the system are:

- Home – The landing page of the website.
- About – This section guides a new visitor with the purpose of analysis and the following sections.
- Current Trends – This section helps find the link between the education and the occupation, and the future scope of the chosen occupation. STEM occupation – This section helps to lay the foundation as to how many people could relate to the occupation based in their degree field, and were science & engineering graduates.
- Demand – The chart shows the various occupations, grouped by job function. The size of the circles depicts the demand in each field. Mean Annual Wage – The analysis depicts the distribution of the highest paying jobs throughout US, and the wage distribution in the STEM sub-categories.
- Our Suggestions – To conclude the analysis, we have provided a point by point evaluation of the pros and cons of choosing a STEM career. Source – This section lists the data source, and the author information.

2.1 Design Approach

Following simple layouts and rough sketches, the final website was created as follows:

1. Data Source – The data from www.bls.gov was collected, and gathered into csv files. [ste13]
2. Data Modeling – The csv files were cleaned using python and some manual pruning was done, to select data for the required occupations, years under consideration, and converting them to a single numerical format for calculation. Map data for US, was collected as a JSON with topographical features.

3. Planning – To fit the data well, and depict meaningful outcomes, brainstorming sessions helped to move from rough sketches to final polished layouts.

4. D3 Layouts – The sole purpose of making more interactive, user friendly visualizations was achieved using D3 visualizations and layouts such as chord chart for showing relationship and movement, circle packing chart for showing the demand over the years for each of the sub-categories of STEM.

5. The View – The tabbed html view with menus for navigation was found to be most suitable to represent the analysis. To avoid loading delays from online scripts, some scripts have been downloaded offline such as d3-tip.js. Bootstrap has been used for setting panels and buttons and enhance outlook. JavaScript Arrays came in handy to model csv data into different formats for various components. CSS has been used intensively, both with

6. Version Control- To keep track of all changes, GIT has been used for version control, maintained over 3 branches for the teammates.

2.2 Development Plan

- To have highest user involvement, care has been taken to incorporate transitions, zooming, mouse over, tooltip events while designing the D3 graphics.

- Coloring for the visualization has been chosen to be impactful and not create any inattentional blindness to the viewer. A wide range of color palates, depths, hues and saturation were compared before setting them for the graphics. [Sim05]

- While rendering the elements for choropleth map to represent changes for different years along with a changing bar chart was found to be an ineffective way, with hard to trace changes going parallelly in the map and the chart. To avoid such attentional delays, choropleth map has been created to highlight the profession with highest mean annual wages, and a corresponding donut chart for the chosen state to account for other careers in the field.

- Scales have been chosen to highlight accuracy and precision. For the chord chart, the length of the chord is directly proportional to the number of people. The axis for the future trends clearly indicate the scales and the growth comparing each of the occupations. It can be observed that though there is growth of 15.19% in health-related occupations from 2016 to 2026, but, on scale with other occupations, the growth is lesser from the current point. [1.015]

- Shapes and sizes of nodes and their leaf-nodes play a significant role in describing the demand for STEM graduates in the circle packing hierarchical layout of D3. Care has been taken to avoid misleading presentation.

3 Aspects of visualization

3.1 Current Trends

To answer the question "How much does your education define, where you end up working" – a chord chart has been created for the purpose, which maps the various educational groups to the occupations based on 17533 respondents in 2016. The thickness of the chords depicts the number of people moving from an educational group to an occupation. Future trends can be seen by moving mouse over the desired occupation.— see Table 1

3.2 STEM occupations

We have tried to answer the question: "How useful is Stem & Engineering education in preparing for the job market?" by presenting three different analysis. The purpose is to highlight relationship between the job and degree held by a person, and the expertise that is required to complete any kind of task requiring knowledge of science and engineering. Bar charts are used in this section to depict occupation analysis data.

Occupation Group	Projected Growth (from 2016 to 2026)
Administrative Staff	0.75%
Crafts	6.27 %
Business Management	9.31%
Basic Occupations	5.86%
Health	15.19%
IT	36.12%
Juridical & Cultural	9.82%
Management functions	9.21%
Teachers	9.29%
Salesmen & Service Providers	3.22%
Caretakers	18.23%
Science & Engineering	11.79%
Other	6.72%

Table 1: Projected Employment growth for an occupational group.

1. **Percentage of Science & Engineering highest degree holders with job in field of highest degree:2016.** - In 2016, 35% of graduates who had earned their highest degree in a S&E field worked in an S&E occupation, and 27% worked in the same broad area as their degree. The percentage is higher for those with more advanced degrees, most notably among doctorate holders. Grouped bar chart was the best way to represent this data.[ste17b]

2. **Relationship between job and degree among workers whose highest degree is in S&E, by type of occupation**: In 2016, about 35% of S&E highest degree holders were employed in occupations classified as S&E, 14% in occupations classified as S&E-related, and 51% in non-S&E occupations.

3. **Job Requirements for technical expertise in S&E**: In 2016, 16.5 million college graduates reported that their jobs required at least a bachelor's degree level of technical expertise in one or more S&E fields. This figure is more than triple the estimated 5.4 million college graduates employed in S&E occupations in 2016, indicating that technical expertise in S&E is important for other types of occupations.

3.3 STEM Employment Demand

STEM occupations are identified in a variety of ways. This section uses a list based on the Standard Occupational Classification Manual to analyze occupations from five groups, including computer and mathematics; architecture and engineering; and life, physical, and social sciences. Across all levels of educational attainment, the largest group of STEM jobs is within the computer and math fields, which account for close to half (49 percent) of all STEM employment. Second are engineering and surveying occupations with 30 percent of all STEM employment, while 12 percent are in the physical and life sciences and 9 percent in STEM management jobs. [ste17a]

To depict this data, we have implemented zoomable circle packing layout of d3.hierarchy because we wanted to show the hierarchy of STEM occupations and at the same time circle size shows the employment demand for the year 2011-2017. We select the year through slider. (See table 2)

3.4 STEM occupations - Mean Annual Wage

The STEM major group comprises of the various occupations like - Computer and Information Research Scientists ; Computer Systems Analysts ;Information Security Analysts ; Computer Programmers ; Software Developers, Applications ; Software Developers, Systems Software ;Web Developers ; Database Administrators ; Network and Computer Systems Administrators ; Computer Network Architects ; Computer User Support Specialists ; Computer Network Support Specialists ; Computer Occupations, All Other ; Actuaries ; Mathematicians ;Operations Research Analysts ; Statisticians ; Mathematical Technicians ; Mathematical Science Occupations.Each occupation consist of mean annual wage for each State in US. The annual wages will help in deciding the occupational growth and demand in each State. (See table 3 and 4).

Occupation	Projected 2016–26 (percent)	Employment 2016	Employment 2026
Information security analysts	37	75,100	102,500
Operations research analysts	27	73,200	92,700
Statisticians	27	27,600	34,900
Biomedical engineers	27	19,400	24,600
Actuaries	26	24,300	30,600
Petroleum engineers	26	38,500	48,400
Computer systems analysts	25	520,600	648,400
Software developers, applications	23	613,000	752,900
Mathematicians	23	3,500	4,300
Software developers, systems software	20	405,000	487,800
Computer user support specialists	20	547,700	658,500
Web developers	20	141,400	169,900
Civil engineers	20	272,900	326,600
Biological science teachers, postsec	20	61,400	73,400
Environmental science technicians	19	32,800	38,900

Table 2: Projected Employment growth for an occupational group.

State	Occ. Title	Mean Annual Wage
Alabama	Computer Information Systems Managers	117,950
Alabama	Computer Systems Analysts	80,230
Alabama	Computer Programmers	83,450
Alabama	Software Developers Applications	93,340
Alabama	Software Developers Systems Software	100,170
Alabama	Network Computer Systems Administrators	71,790
Alabama	Computer User Support Specialists	46,870
Alabama	Civil Engineers	78,100
Alabama	Mechanical Engineers	88,600
Alabama	Sales Representatives Technical Scientific	86,290

Table 3: The employment estimate data base and mean annual wages is described below

To represent this data, we have shown chloropleth map of United States, where darker shade corresponds to state with the higher wage. For interactivity we have added 2 things :

- On hovering over a state, will add a tooltip, which shows the profession with highest wage for that state.
- Also simultaneously, on hovering over the state a pie chart showing the distribution of occupation and wages is created on the right side of map.

4 Evaluation

4.1 Verification Checklist

We have followed the below checklist while evaluating our visualization model:

1. Accuracy – while drawing different charts, presenting misleading information is quite common. To avoid the situation, care has been taken to map all elements in a visualization to a single scale and dimension.

Percentile	10%	25%	50% (Median)	75%	90%
Annual Wage (2)	$42,650	$59,120	$82,830	$111,690	$141,250

Table 4: The employment estimate data base and mean annual wages is described below

2. Hierarchical structure – It has been ensured that the text sizes for titles, sub-titles and text is following a hierarchy and is readable at the same time.

3. Data Labeling – Information related to a chart has been provided alongside like legends, important findings to bring a clearer picture to the reader.

4. Formatting – The text fonts, and coloring properties are kept uniform throughout, and defined through a single css file.

5. Cluttering – To avoid cluttering, data has been abbreviated and present as a legend or tooltip text, to enhance readability and understanding of the idea under focus.

6. Cairo's visualization wheel – The model is built in consideration of the Cairo's Visualization Wheel Principles. The model has multidimensionality, functionality and are dense. In order to represent multiple features in single visualization, the model is built interactive. The charts were built with the ability to define various attributes in a single visualization.

4.2 Feedback from Critique Group

There were multiple sessions and SLACK conversations held with the Critique Group in order to get constant feedback which helped directly in evaluating the model every time. The feedback, suggestions were considered and taken into account every time. The suggestions provided by the critique group further helped us in the evaluation process from the perspective of an uninformed reader.

5 Future Work

1. Adding Future Predictions to Demand data by applying machine learning on the past data that will give us predictions.

2. Sorting the pie chart on the basis of legend.

3. Resizing of charts as per window size

References

[1.015] mbostock 1.0.7. D3 scales. *https://github.com/d3/d3-scale*, 2015.

[Sim05] Robert Simmon. Coloring in visualization. *https://earthobservatory.nasa.gov*, 2005.

[ste13] You're a what? ornithologist" in the summer 2013 issue. *www.bls.gov/ooq/2013/summer/yawhat.pdf*, 2013.

[ste14] Stem 101: Intro to tomorrow's jobs. *www.bls.gov/careeroutlook*, 2014.

[ste17a] Stem jobs: 2017 update, by u.s. department of commerce economics and statistics administration office of the chief economist. *http://www.esa.doc.gov/sites/default/files/stem-jobs-2017-update.pdf*, 2017.

[ste17b] Stem occupations: past, present and future. *https://www.bls.gov/spotlight/2017/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future/pdf/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future.pdf*, 2017.