Grizzly Insights Final Report

Dr. Anca

ITEC 4230

May 3, 2023

Project Description

This project aims to help GGC students choose courses that align with their career goals. We improved the Grizzly Path website using data collected from job sites. We identified in-demand skills with a goal to match them to GGC courses. This will enable students to make more informed decisions about their academic pursuits. By integrating data from job sites into the Grizzly Path website, we have provided GGC students with a powerful tool to help them choose courses that align with their career goals and the skills that are in-demand in the industry.

Team Grizzly Insights

Anel Coralic: Data Analyzer/Team Manager

Sam Downs: Data Modeler/Project Documenter

Ashley Mendez: Data Visualizer/Client Liaison

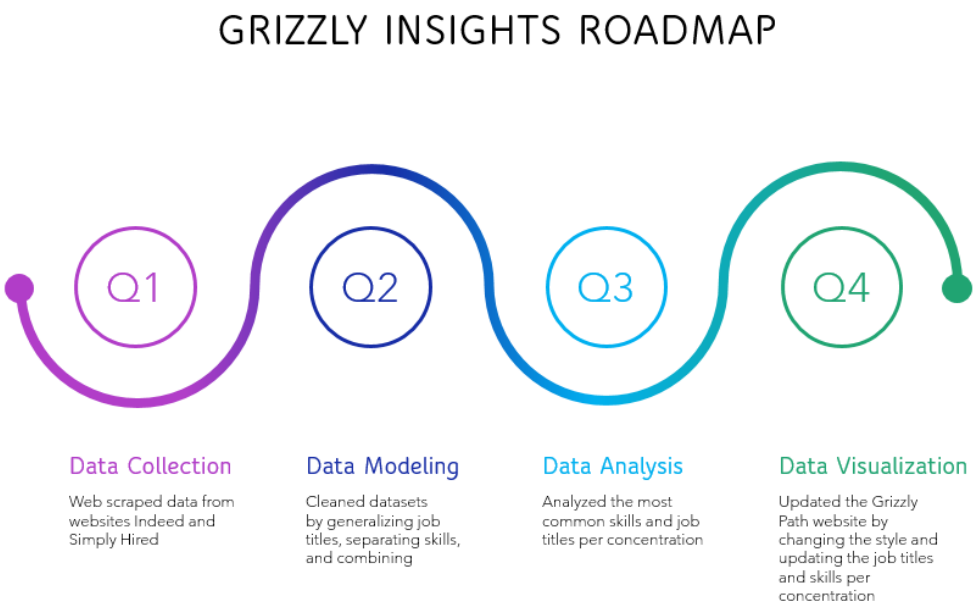
Client

Our client is Dr. Anca Doloc-Mihu. She is an associate professor of information technology at Georgia Gwinnett College. She educates college students about Data Science tools and technologies

Technologies

We used Jira to stay organized and manage our project. We communicated among ourselves and our client with Microsoft Teams, Discord, and Outlook. We used Web Scraper to collect data. We used Python with Jupyter Notebook and Google Colab for our data cleaning and analysis processes. We used HTML, JavaScript, and CSS to edit the Grizzly Paths web application.

Project Flowchart



Data Collection

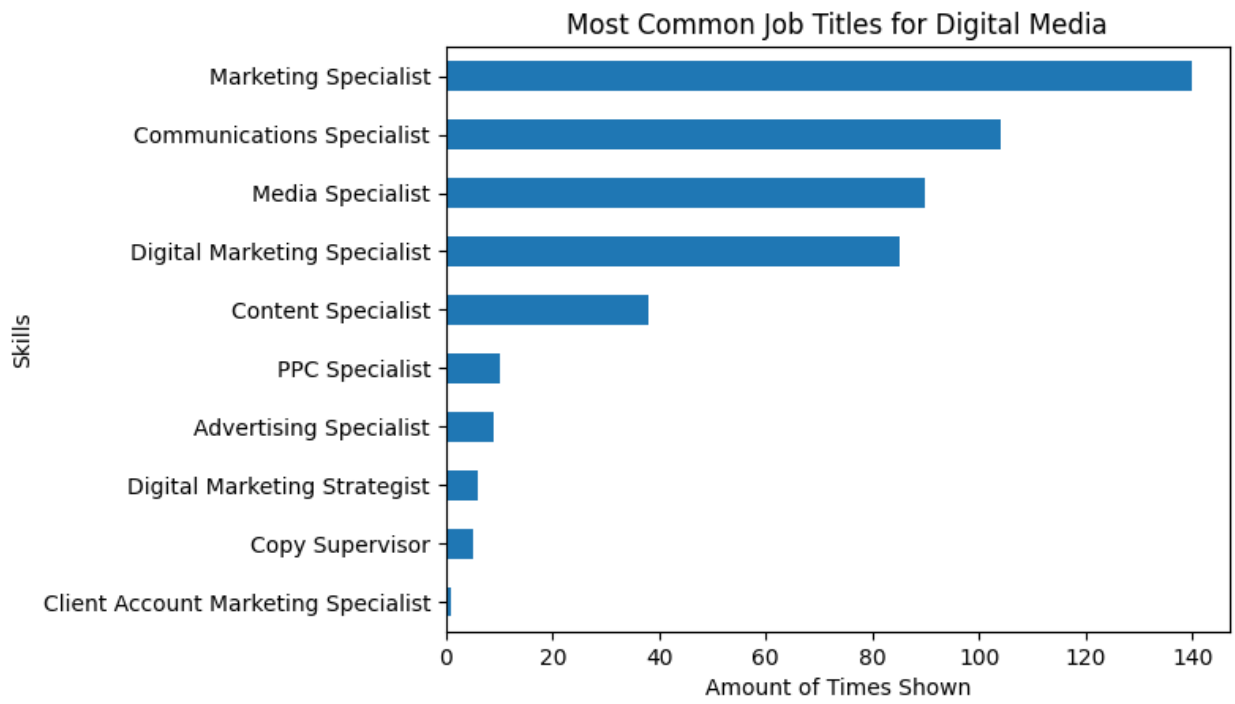
Job listings were scraped from two popular websites: Indeed.com and SimplyHired.com. The information technology concentration we scraped for were Digital Media, Enterprise Systems, Software Development, and Systems and Security. Data was scraped by using a free google chrome extension called Web Scraper. This tool is used by first creating a Sitemap, setting up the selectors to define how the site will be navigated and extracted, and launching the Sitemap. The data is exported to an excel file. The result was about 500 jobs per concentration.

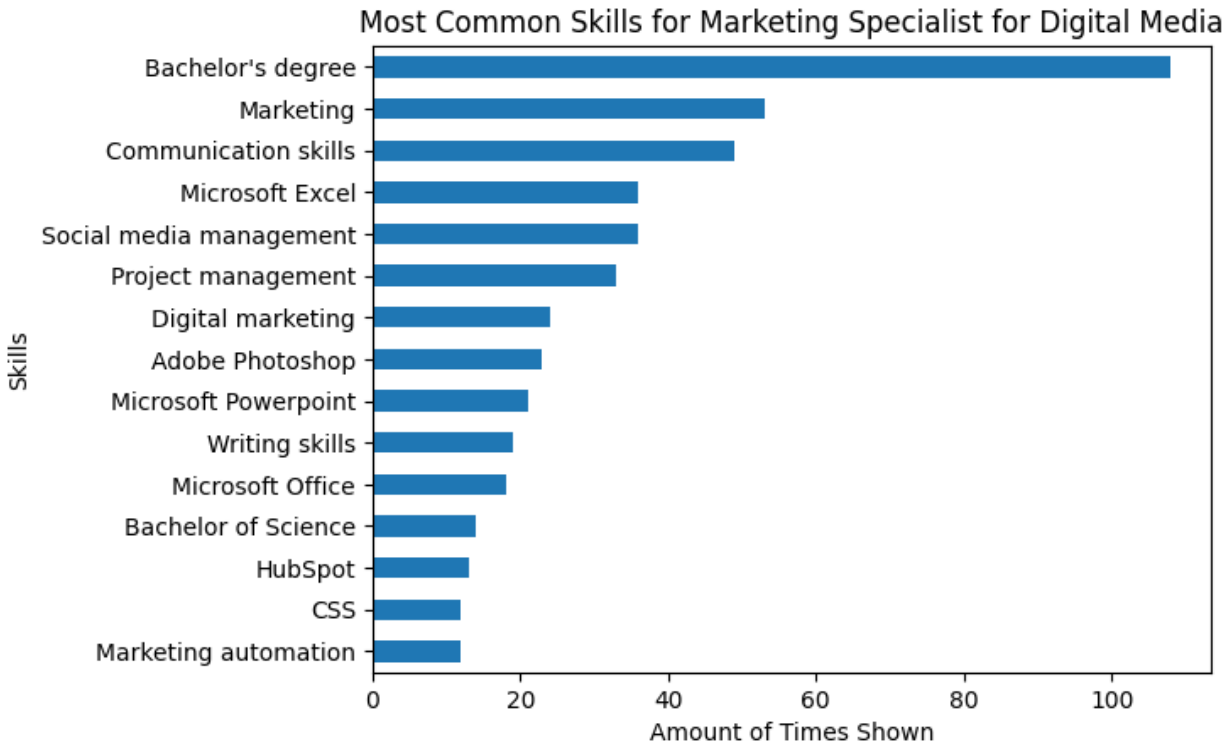
Web Scraper link:
https://chrome.google.com/webstore/detail/web-scraper-free-web-scra/jnhgnonknehpejjnehehllkliplmbmhn?hl=en

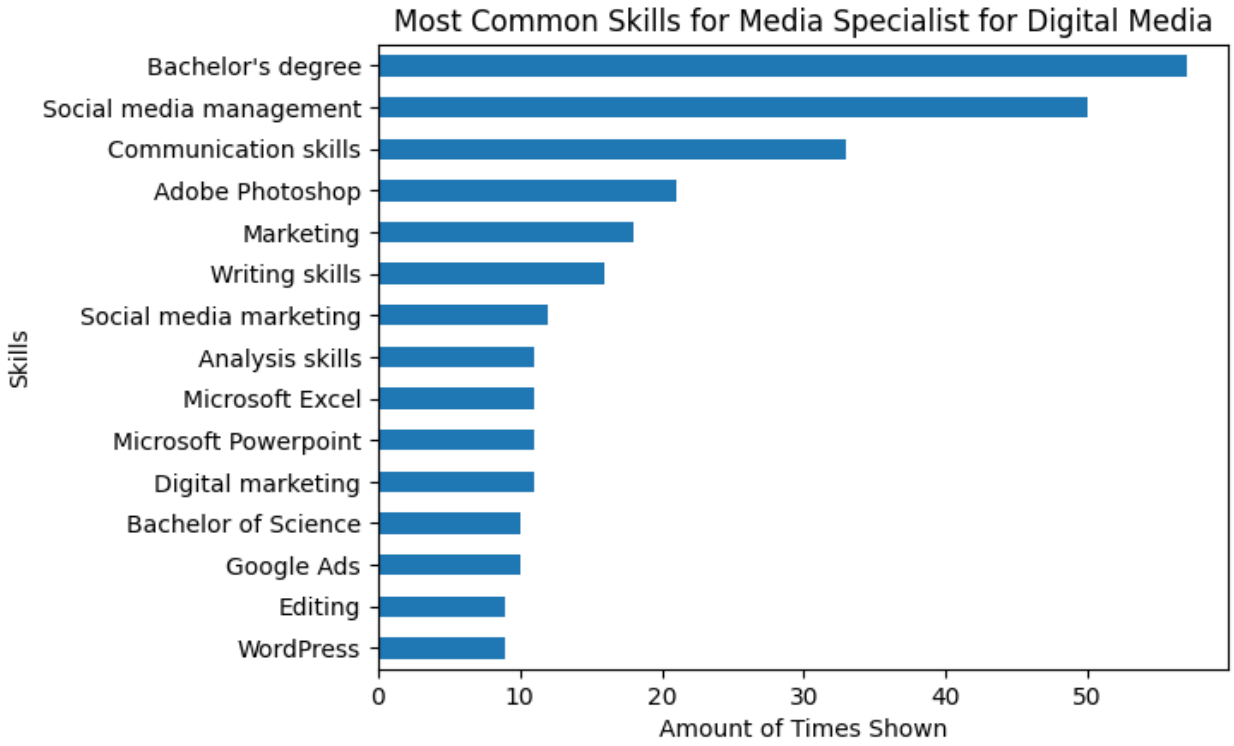
Data Analysis

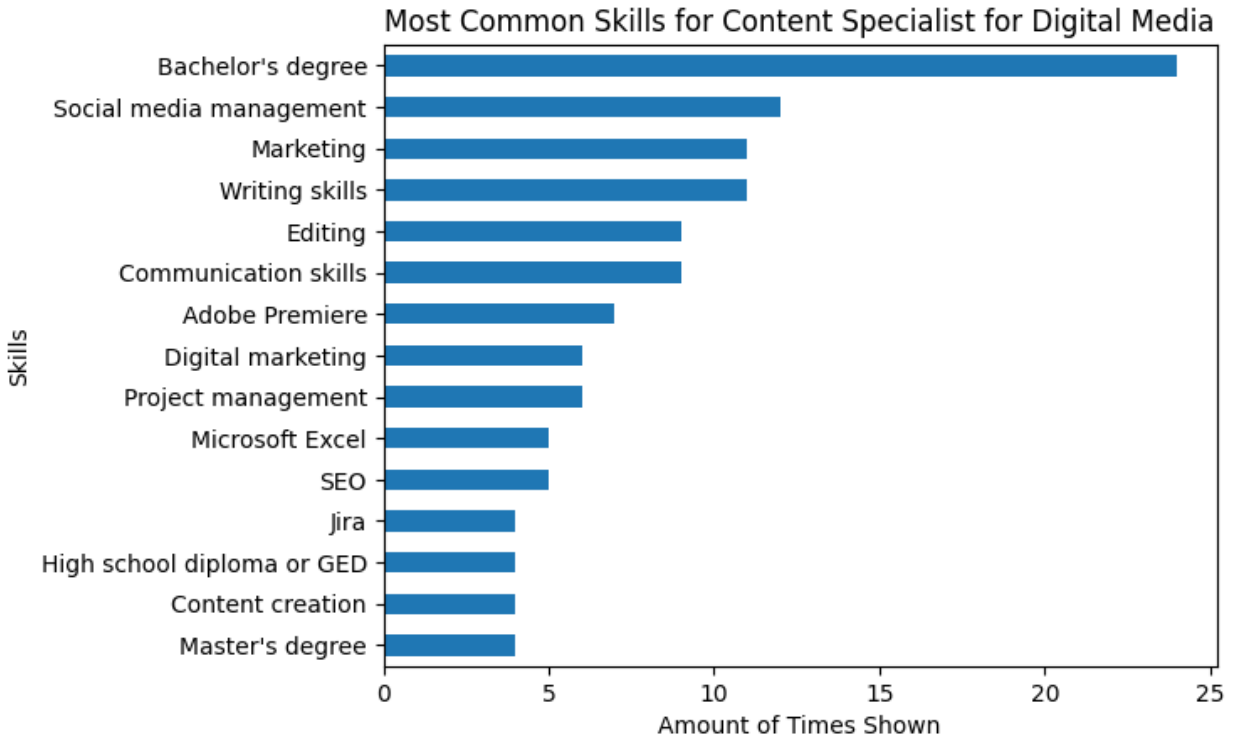
We performed data analysis to identify the most common job titles for each IT concentration. We then determined the most common skills for each of these job titles.To achieve this, we used a method to parse each element of the list in the skills column. Below are our results represented in plots:

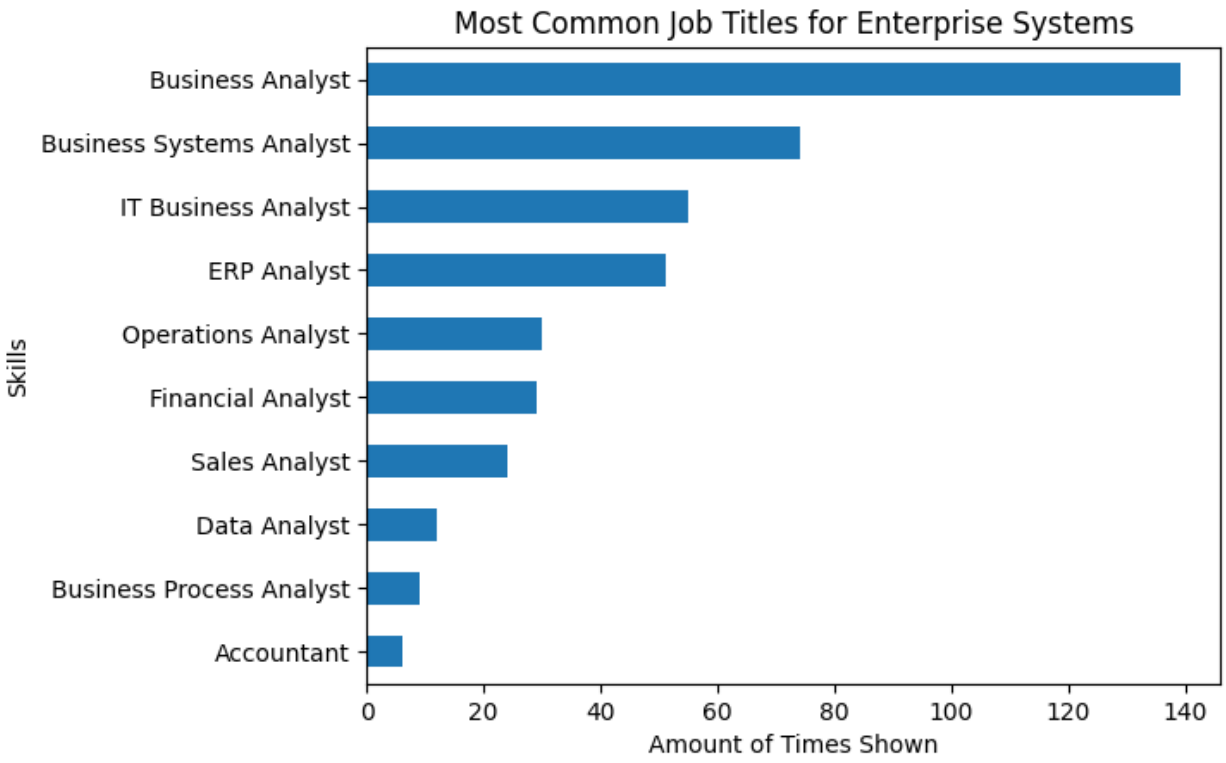
Digital Media

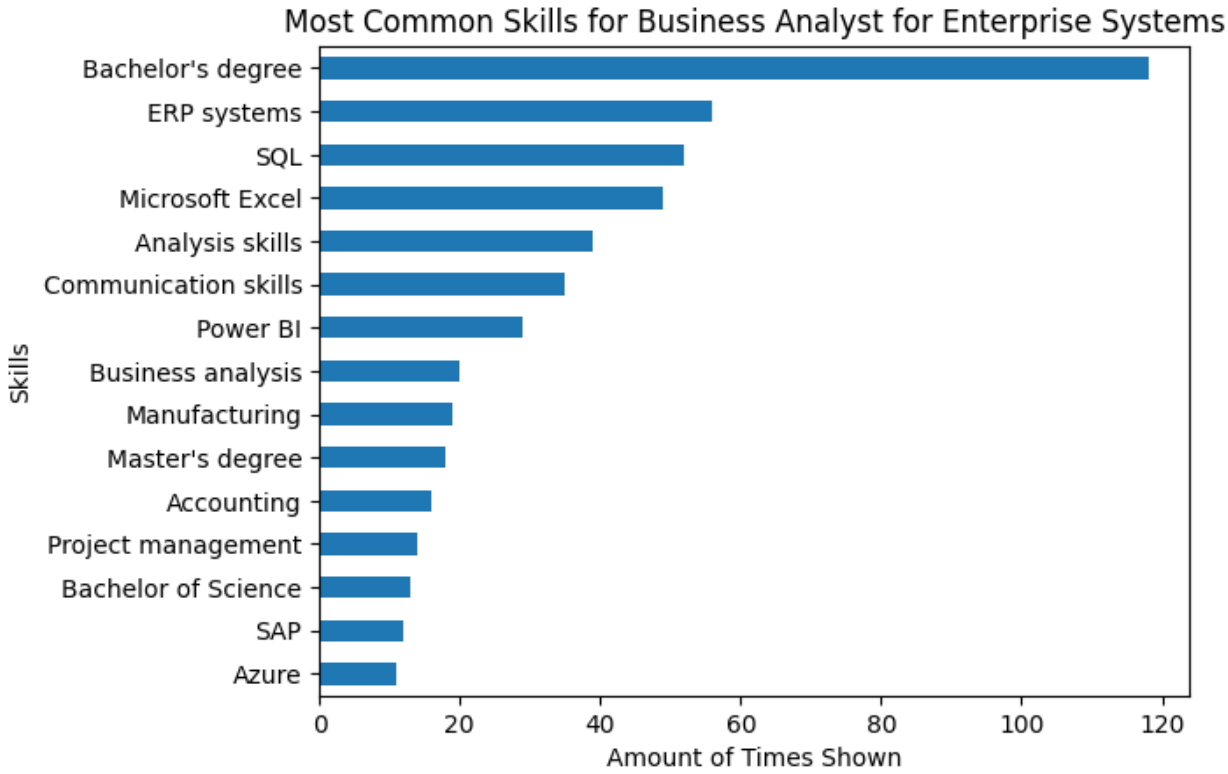

Most Common Job Titles for Digital Media
Marketing Specialist
Communications Specialist
Media Specialist
Digital Marketing Specialist
Content Specialist
PPC Specialist
Advertising Specialist
Digital Marketing Strategist
Copy Supervisor
Client Account Marketing Specialist
Skills
0
20
40
60
80
100
120
140
Amount of Times Shown

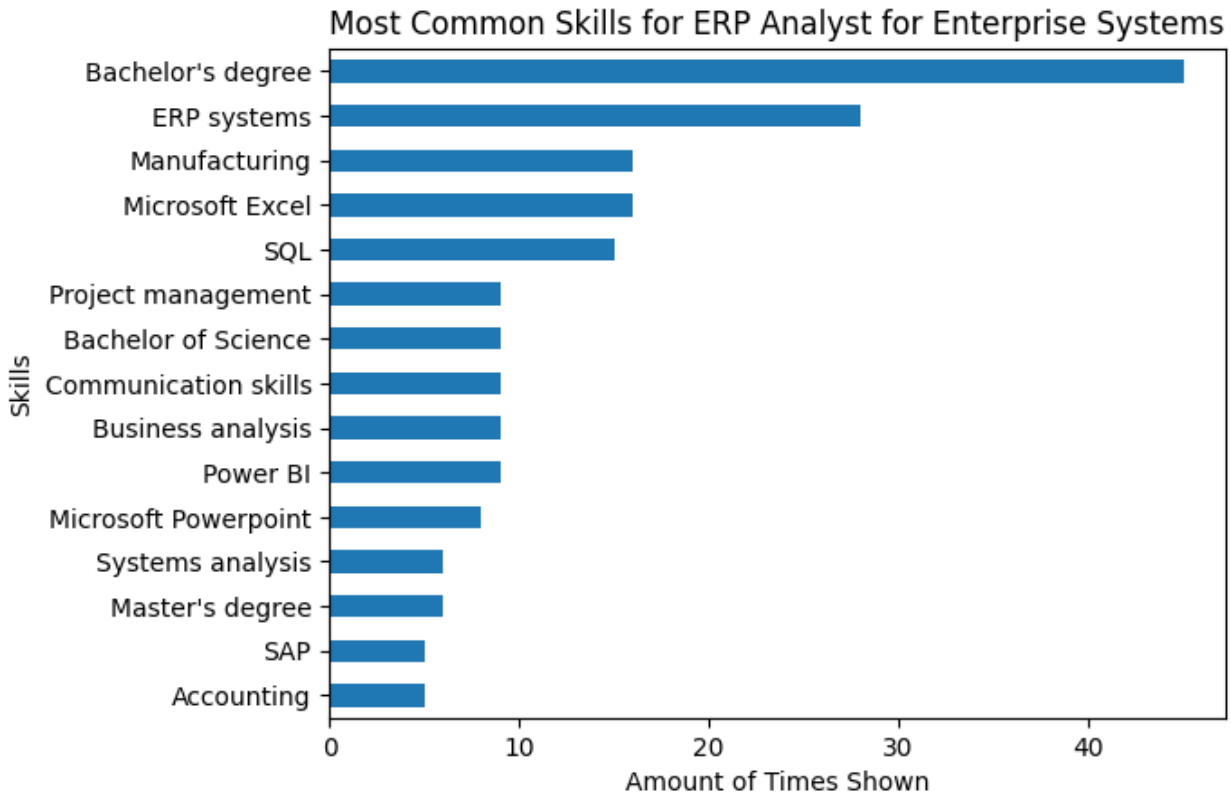

Most Common Skills for Marketing Specialist for Digital Media
Bachelor's degree
Marketing
Communication skills
Microsoft Excel
Social media management
Project management
Digital marketing
Adobe Photoshop
Microsoft Powerpoint
Writing skills
Microsoft Office
Bachelor of Science
HubSpot
CSS
Marketing automation
Skills
0
20
40
60
80
100
Amount of Times Shown

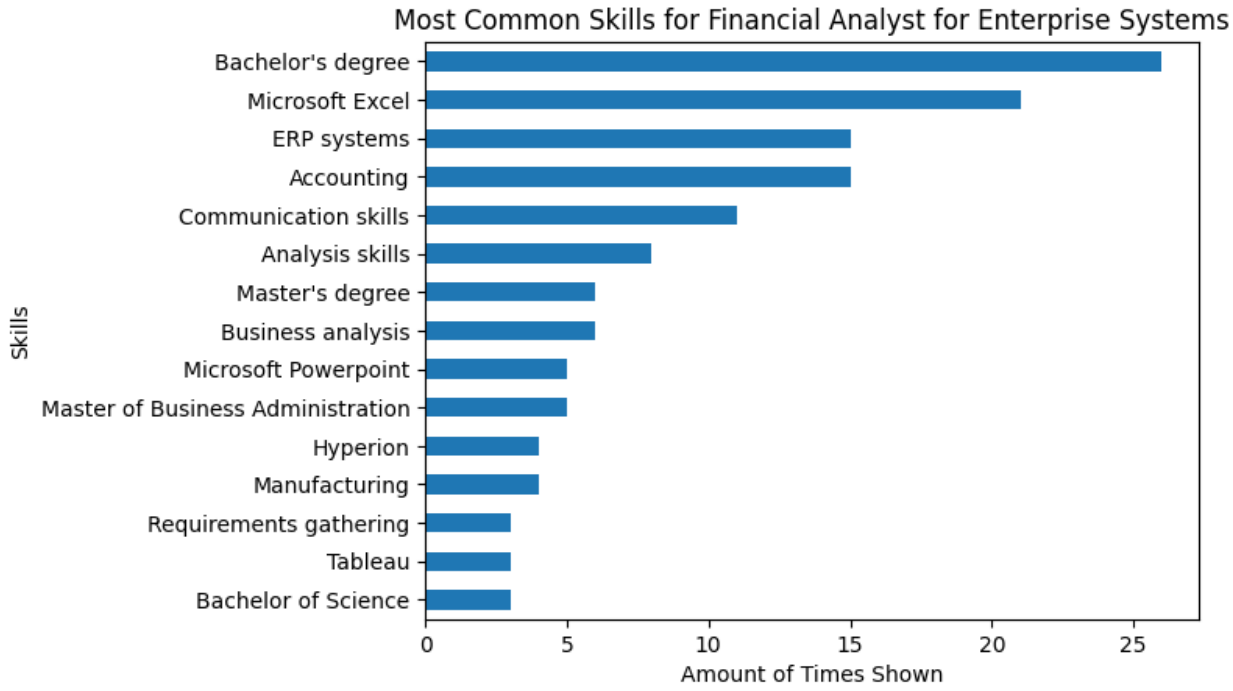

Most Common Skills for Media Specialist for Digital Media
Bachelor's degree
Social media management
Communication skills
Adobe Photoshop
Marketing
Writing skills
Social media marketing
Analysis skills
Microsoft Excel
Microsoft Powerpoint
Digital marketing
Bachelor of Science
Google Ads
Editing
WordPress
Skills
0
10
20
30
40
50
Amount of Times Shown


Most Common Skills for Content Specialist for Digital Media
Bachelor's degree
Social media management
Marketing
Writing skills
Editing
Communication skills
Adobe Premiere
Digital marketing
Project management
Microsoft Excel
SEO
Jira
High school diploma or GED
Content creation
Master's degree
Skills
0
5
10
15
20
25
Amount of Times Shown

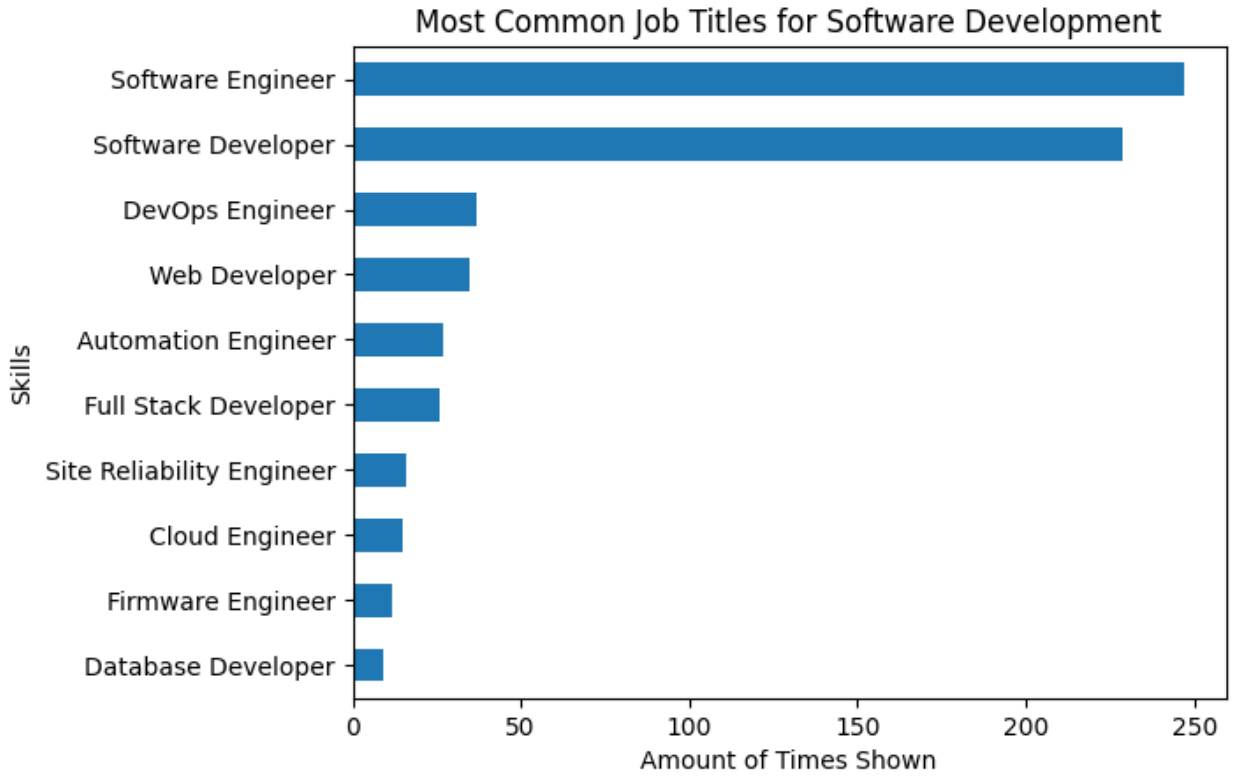

Most Common Job Titles for Enterprise Systems
Business Analyst
Business Systems Analyst
IT Business Analyst
ERP Analyst
Operations Analyst
Financial Analyst
Sales Analyst
Data Analyst
Business Process Analyst
Accountant
Skills
0
20
40
60
80
100
120
140
Amount of Times Shown

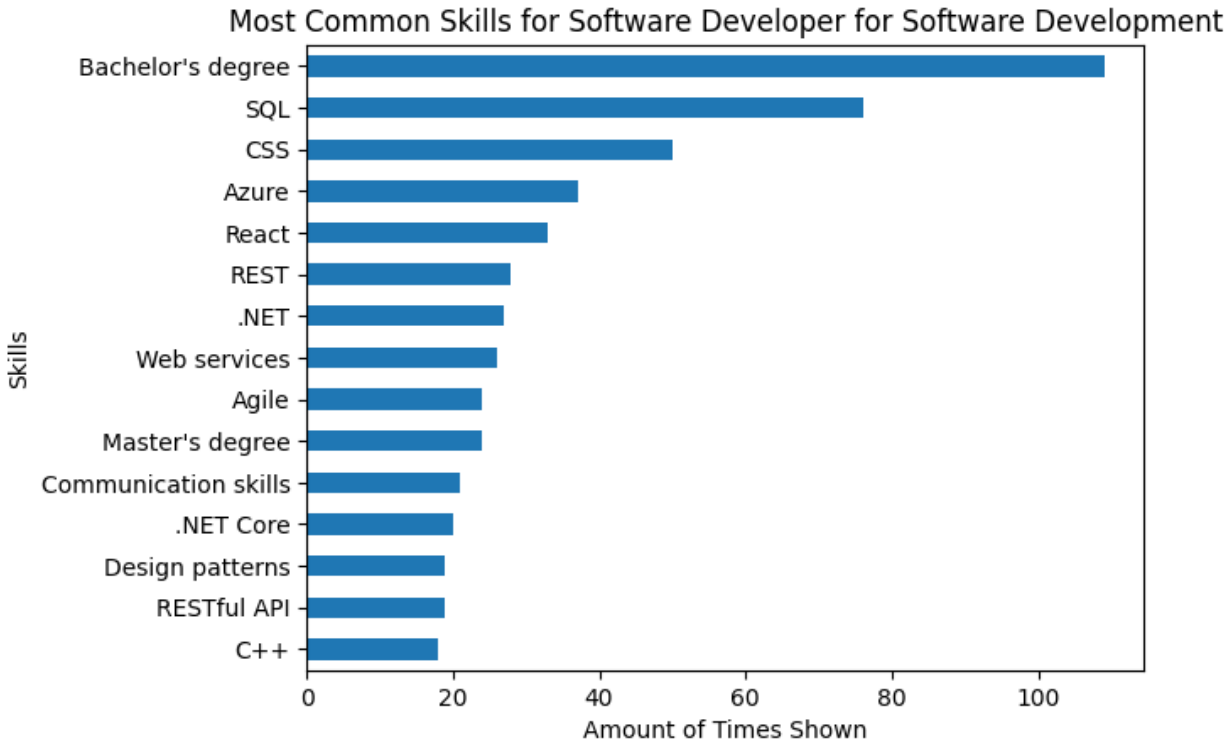

Most Common Skills for Business Analyst for Enterprise Systems
Bachelor's degree
ERP systems
SQL
Microsoft Excel
Analysis skills
Communication skills
Power BI
Business analysis
Manufacturing
Master's degree
Accounting
Project management
Bachelor of Science
SAP
Azure
Skills
0
20
40
60
80
100
120
Amount of Times Shown

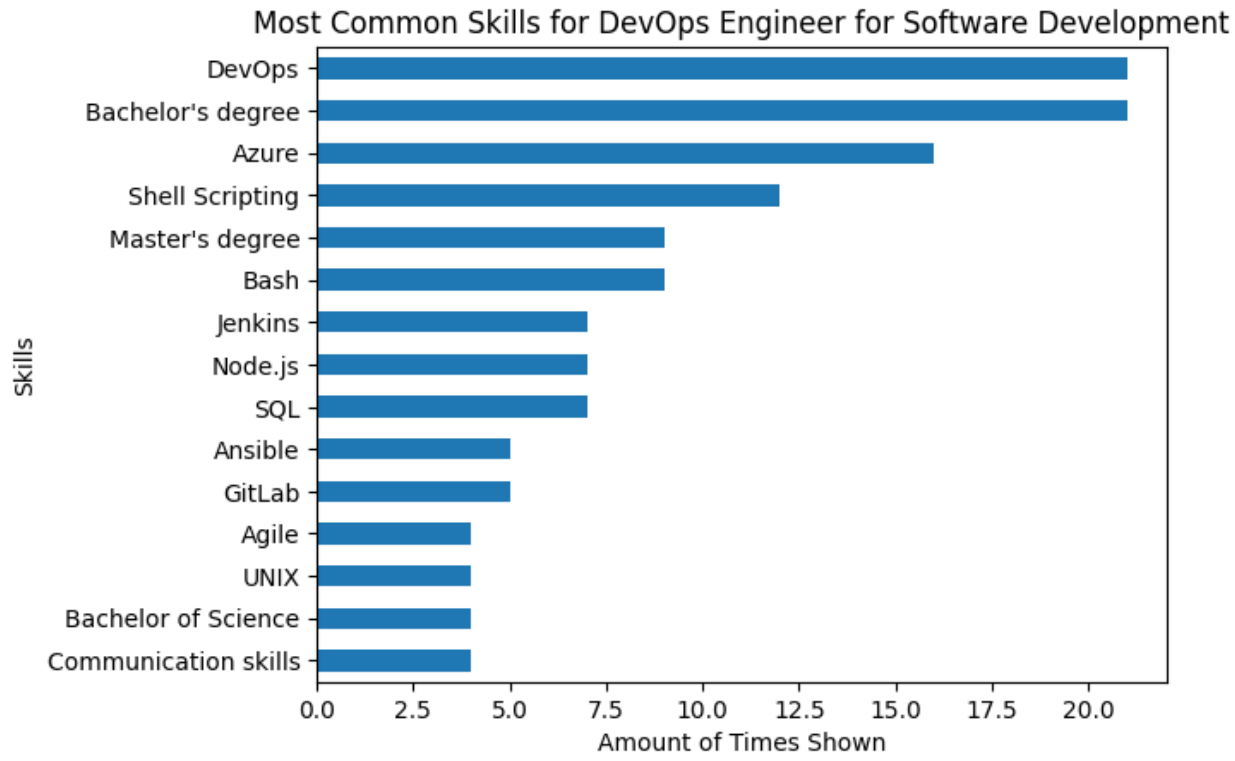


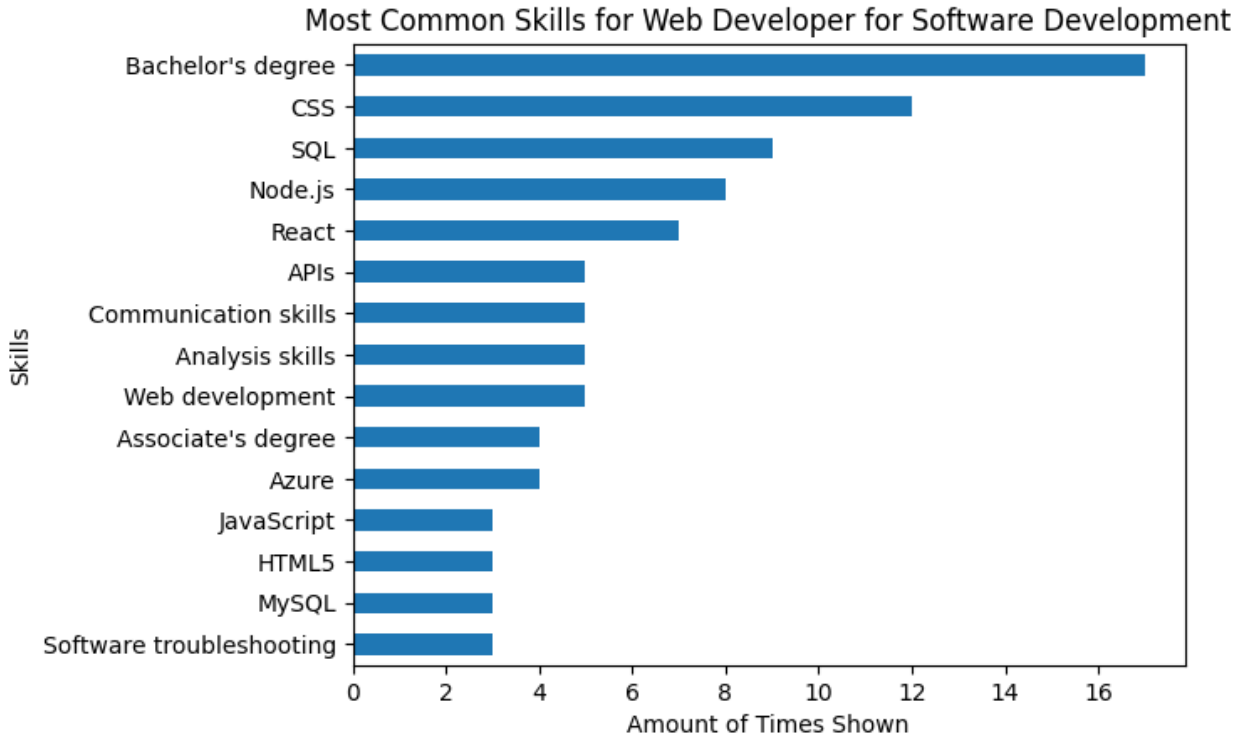


Software Development

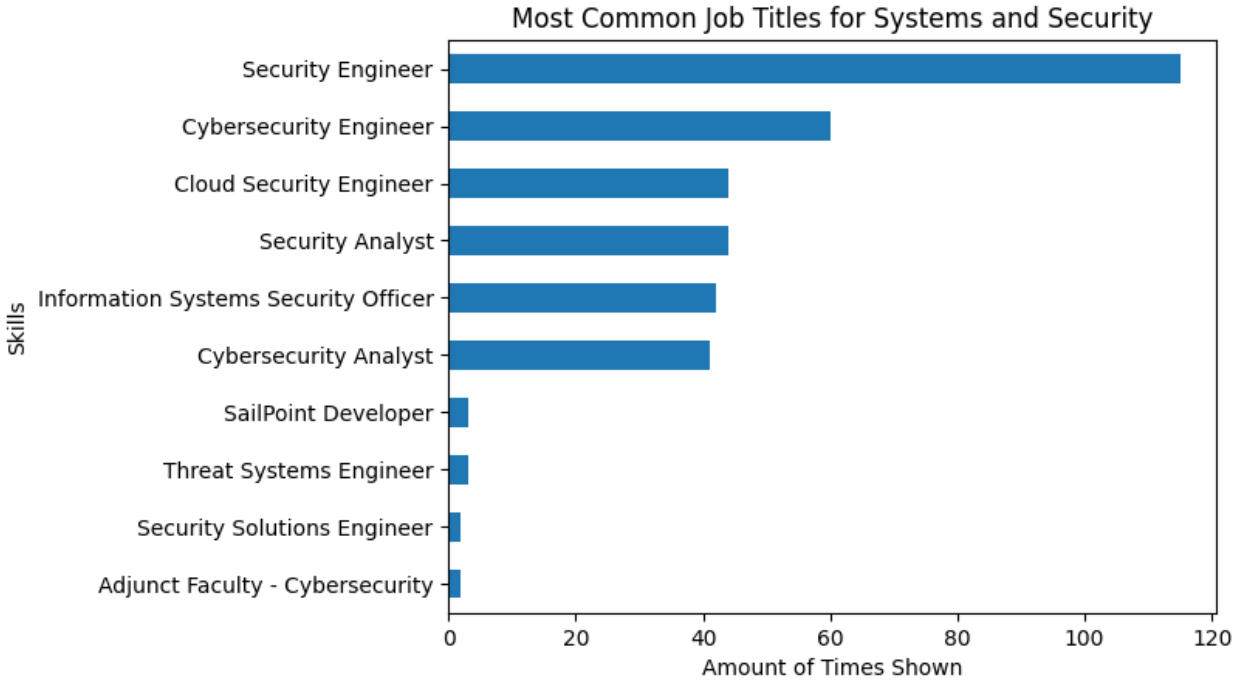

Most Common Job Titles for Software Development
Software Engineer
Software Developer
DevOps Engineer
Web Developer
Automation Engineer
Full Stack Developer
Site Reliability Engineer
Cloud Engineer
Firmware Engineer
Database Developer
Skills
0
50
100
150
200
250
Amount of Times Shown

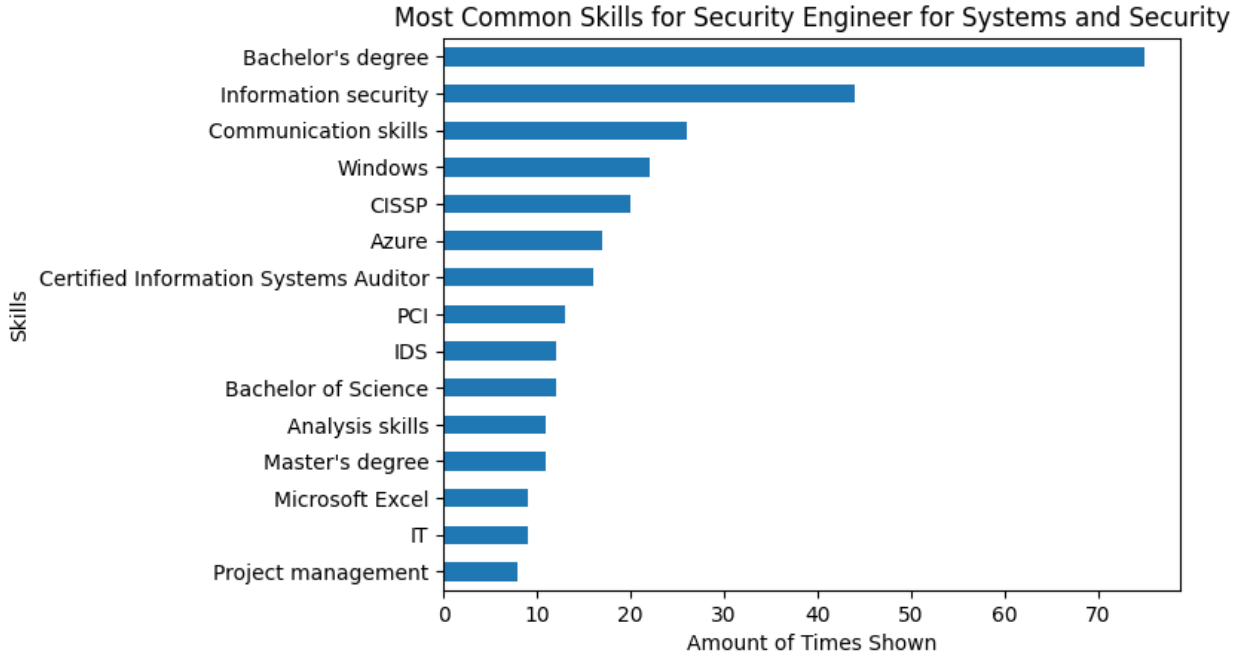

Most Common Skills for Software Developer for Software Development
Bachelor's degree
SQL
CSS
Azure
React
REST
.NET
Web services
Agile
Master's degree
Communication skills
.NET Core
Design patterns
RESTful API
C++
Skills
0
20
40
60
80
100
Amount of Times Shown

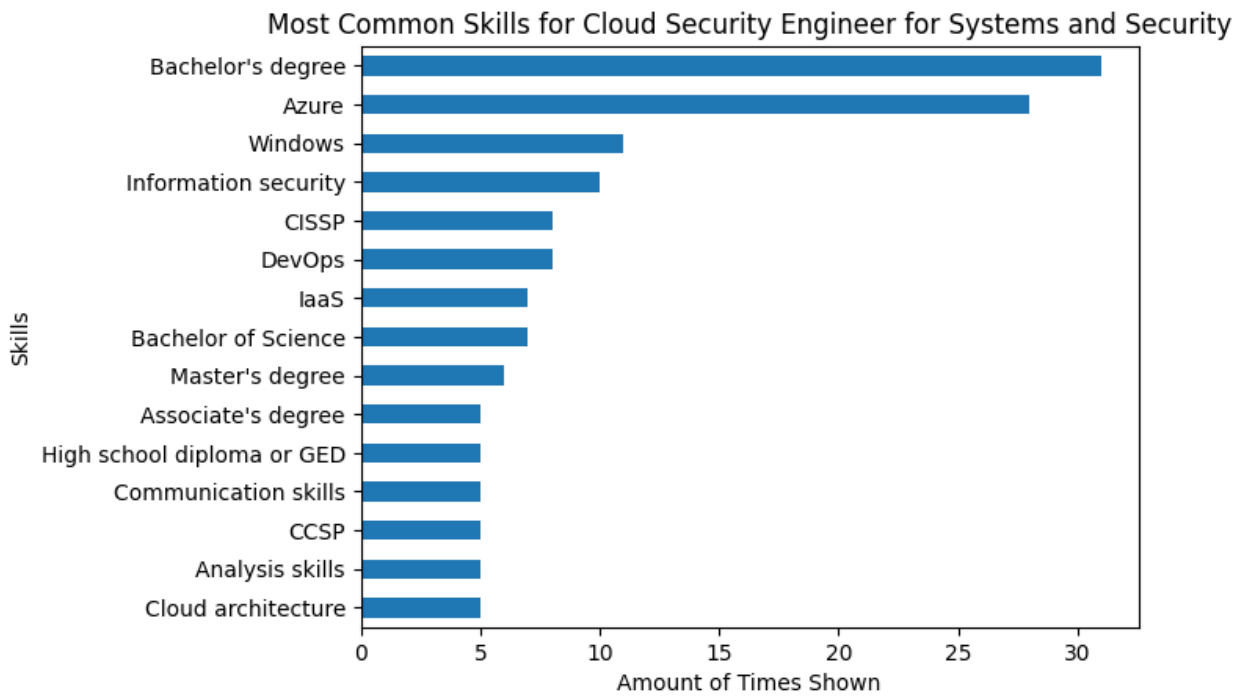

Most Common Skills for DevOps Engineer for Software Development
DevOps
Bachelor's degree
Azure
Shell Scripting
Master's degree
Bash
Jenkins
Node.js
SQL
Ansible
GitLab
Agile
UNIX
Bachelor of Science
Communication skills
Skills
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Amount of Times Shown

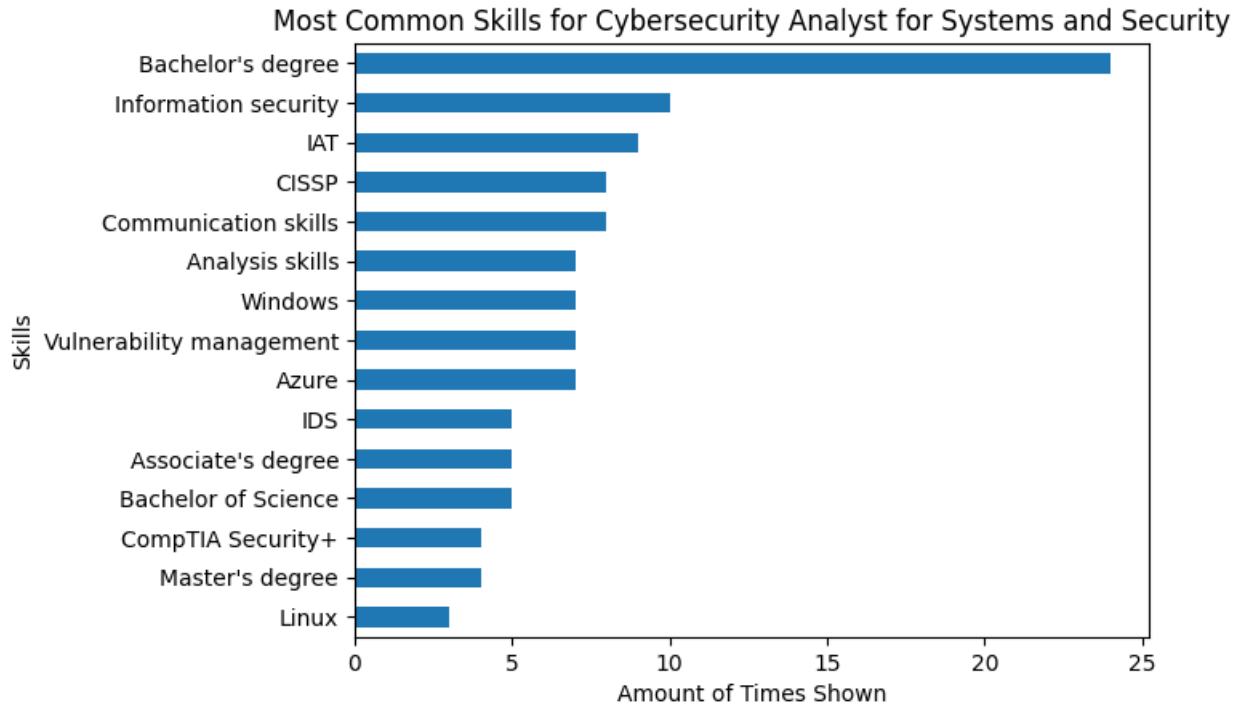

Most Common Skills for Web Developer for Software Development
Bachelor's degree
CSS
SQL
Node.js
React
APIs
Communication skills
Analysis skills
Web development
Associate's degree
Azure
JavaScript
HTML5
MySQL
Software troubleshooting
Skills
0
2
4
6
8
10
12
14
16
Amount of Times Shown

Systems and Security









Three Iterations

For the first iteration, we collected data from job search websites, specifically Indeed and Simply Hired, by utilizing a Web Scraper extension. For the second Iteration, the datasets were cleaned by performing generalization, separating skills, and combining them. For the third Iteration, we analyzed and visualized the data using the Grizzly Path website.

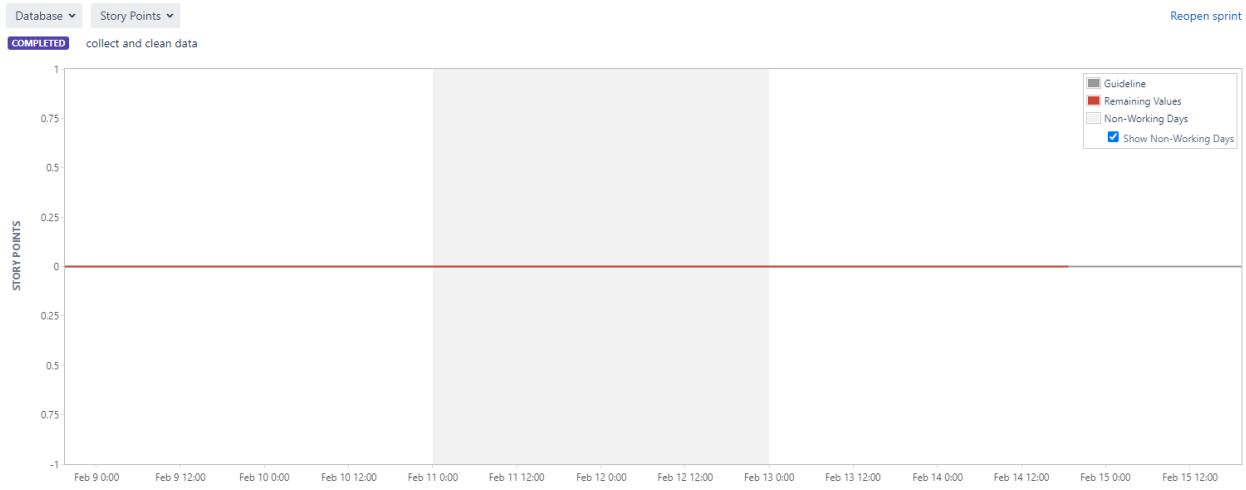
Github Repository

Our Github is shared among teams that worked on this project prior to us. Our team is the Spring 2023 team. You can find our work in the "SPR2023" folder. There you can find our data files and notebook.

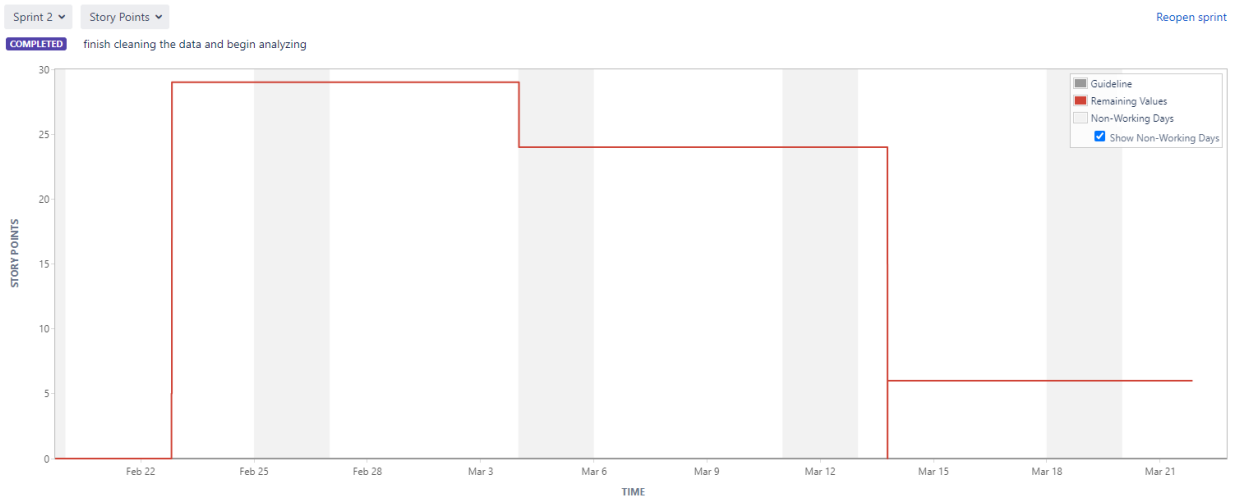
Github repository link: https://github.com/GGC-DSA/itskills/tree/main/SPR2023
Grizzly Paths web application link: https://ggc-dsa.github.io/itskills/

Jira Dashboard

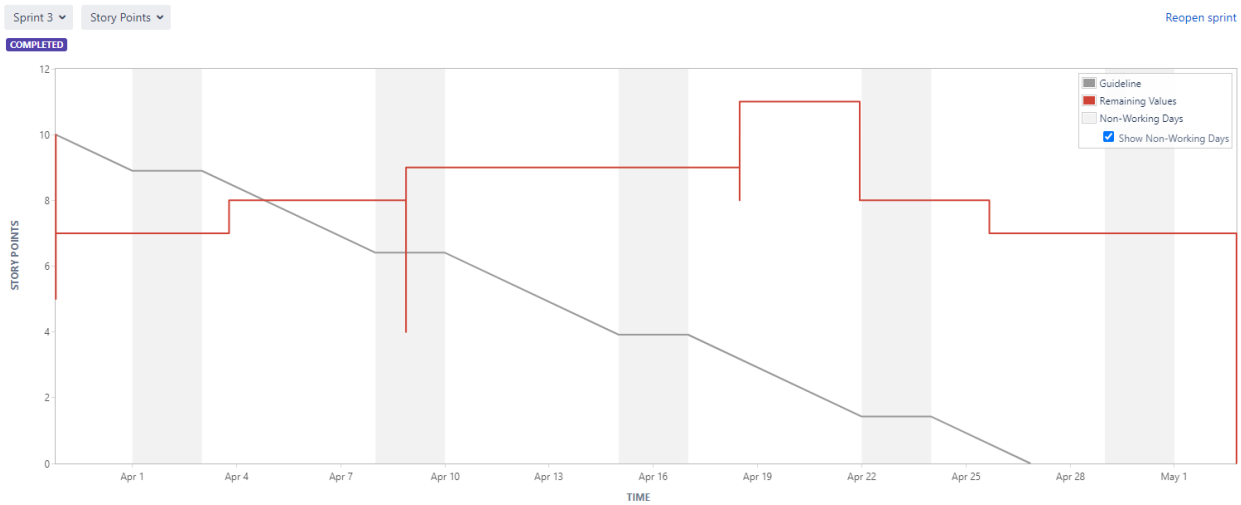
Sprint 1 Burndown Chart



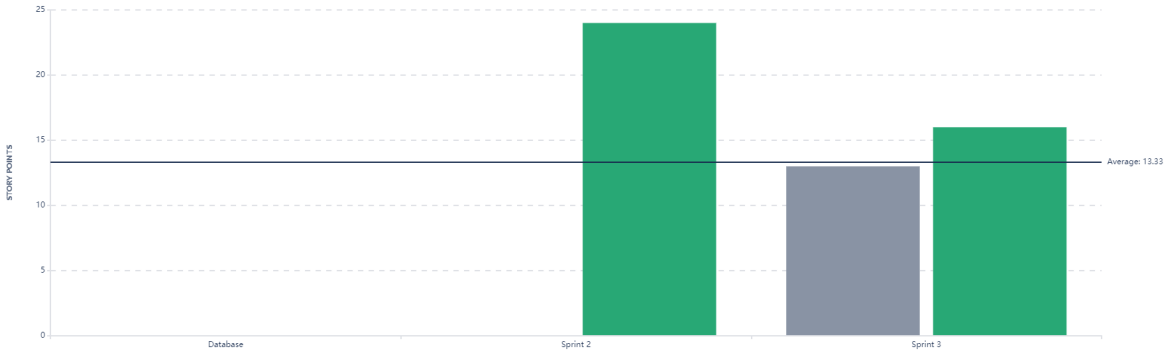
Sprint 2 Burndown Chart



Sprint 3 Burndown Chart



Velocity Chart



Screencast

Youtube link: https://youtu.be/NyiuNdNvz9k

Features Implemented

We implemented a home page for the web application. On the home page, we added a section that includes the names of all who have worked on the web application. We also added a drop down menu where users can select their desired information technology concentration, the page for that concentration is loaded, and a visualization is formed for that specific concentration.

Known Issues

The project needs more data to be collected to represent more variety in the different jobs in each concentration better. Data needs to be collected from student surveys to understand what skills are learned in which courses, so that the courses at GGC can be mapped to jobs. The visualization shows the five most common skills for each job title, but it would be helpful for students to know all of the recommended skills for each position. Analysis needs to be done on the three different categories of skills rather than looking at the skills as a whole, and this should be represented somehow on the web application.

To Do List

1) Create/Update GGC class survey. Needs to be user friendly and easier to extract data.
2) Ask GGC IT students to complete survey
3) Associate skills from job titles to GGC courses.
4) Update Grizzly Path website with GGC courses