ALY 6080:
INTEGRATED EXPERIENTIAL LEARNING

Assignment 6: Group Project- CoverQuick Datasets Roadmap

Submitted To:
Dr. Chinthaka Pathum Dinesh, PhD, Prof. Herath Gedara,
Faculty Lecturer

Submitted By:
Abhilash Dikshit
Kush Patel
Siddharth Alashi

Academic Term: Spring 2023
Graduate Students at Northeastern University, Vancouver, BC,
Canada
Master of Professional Studies in Analytics

May 24, 2023

I.   Abstract:

This roadmap outlines the plan for our analytic project, detailing the approach, milestones, job assignments, key risks, mitigation strategies, measures of success, and presentation method. The project aims to answer specific requirements by employing various models, tools, and techniques. This roadmap serves as a guide to ensure progress, collaboration, risk mitigation, and successful delivery of the proof of concept.

II.   Introduction:

In today's competitive job market, employers are looking for ways to identify the most qualified candidates for job openings. One way to accomplish this is to create a database of job applicants, which can provide insights into their education, work experience, and skills. However, before building such a database, it is necessary to prepare the data. Data preparation involves cleaning, transforming, and exploring the data to ensure it is accurate, consistent, and meaningful.

1. Analytic Approach:
To address the project requirements, we will adopt the following analytic approach:

A. Data Preprocessing: Clean and preprocess the dataset, handling missing values, outliers, and data quality issues. Utilize techniques such as data imputation, feature scaling, and handling categorical variables.

CoverQuick No Job Description Dataset:
==

Display Raw Dataset:

	id	conte
0	clezbskz2000ryi2yj9sgidax	{"awards": {"awards": []}, "header": {"role":
1	clezeac8k001myg2u6yph9yz0	{"awards": {"awards": []}, "header": {"role":
2	clezh9pxu0185yp2yo6ztscpf	{"awards": {"awards": []}, "header": {"role":
3	clf06bf5b014dx22ur1bph3sd	{"awards": {"awards": []}, "header": {"role":
4	clezdavhy000eyg2uh55ycnb4	{"awards": {"awards": []}, "header": {"role":
...	...	
13193	clbmqeb0p00bowy50puy0zh47	{"awards": {"awards": []}, "header": {"role":
13194	cld9mmcjj004mx32l3q1bn6ju	{"awards": {"awards": []}, "header": {"role":
13195	cld8csfqd002yyc2m1l7vqsuv	{"awards": {"awards": []}, "header": {"role":
13196	clelqs1nd0002x12vg9vjumyg	{"awards": {"awards": []}, "header": {"role":
13197	cldy3pez8000xxd2xvvj0fpyk	{"awards": {"awards": []}, "header": {"role":

13198 rows × 2 columns

Display datatypes of respective columns in dataset:

	id	content
0	object	object

Showing max, min length and NA values:

Column	Max Length	Min Length	NA Count
id	25	25	0
content	30743	484	0

DATASET SHAPE:
==
(13198, 2)

B. Exploratory Data Analysis (EDA): Conduct comprehensive EDA to gain insights into the dataset, identify patterns, correlations, and anomalies. Visualize the data using graphs, charts, and statistical measures.

C. Research Questions:

1. What are the three industries that the majority of CoverQuick's users have applied (No Job Description Dataset) and will apply to (With Job Description Dataset)?
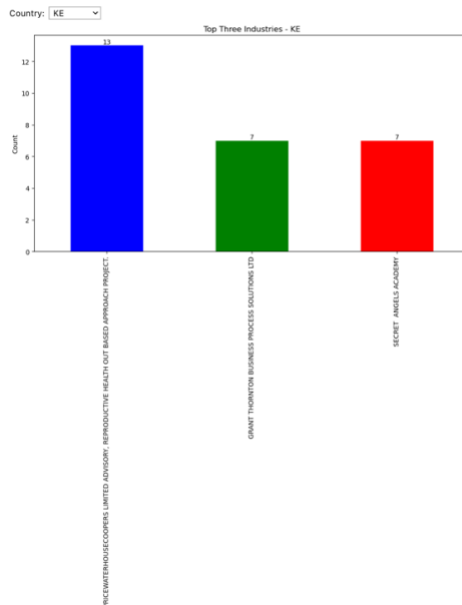
2. Discover trends in demographics and find which industries yield the best and the worst resumes (CoverQuick provides metrics for defining a "Good" resume).
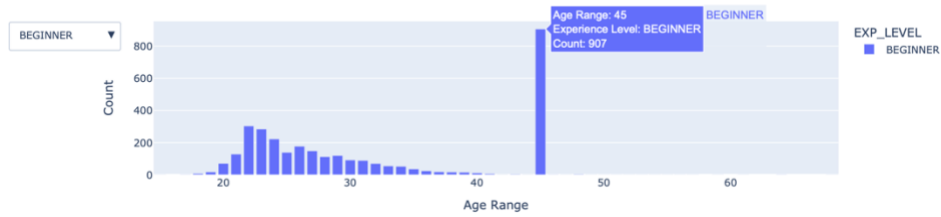

Distribution of Country Codes

Number of Candidates per City in KE

Number of Candidates per City in US
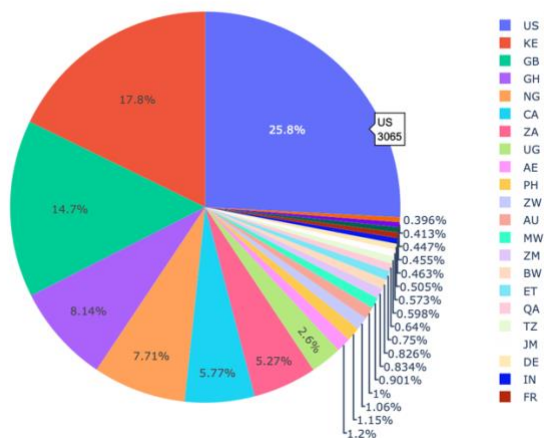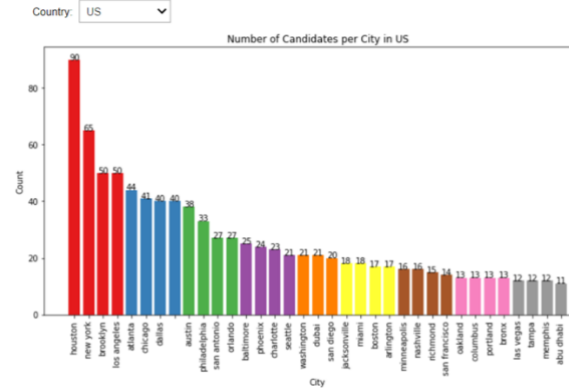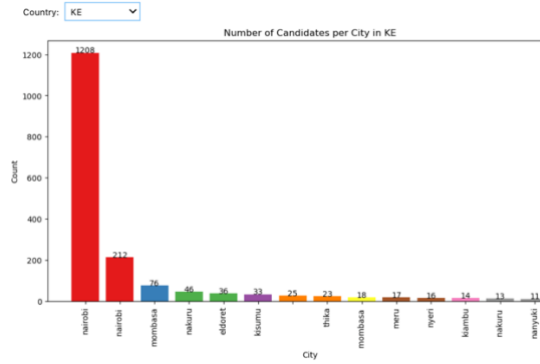
2. Determine the expected age and approximate experience level.

Trend Analysis: Age Range and Experience Level - BEGINNER
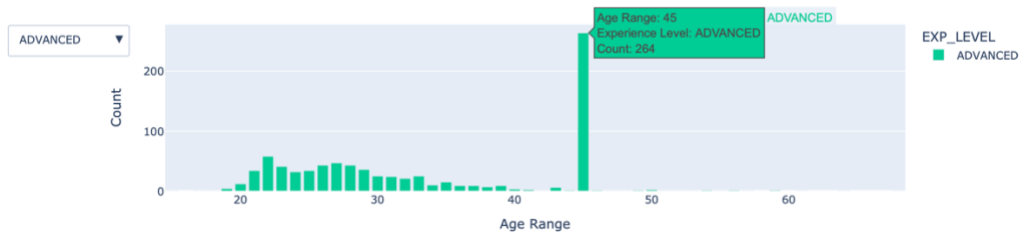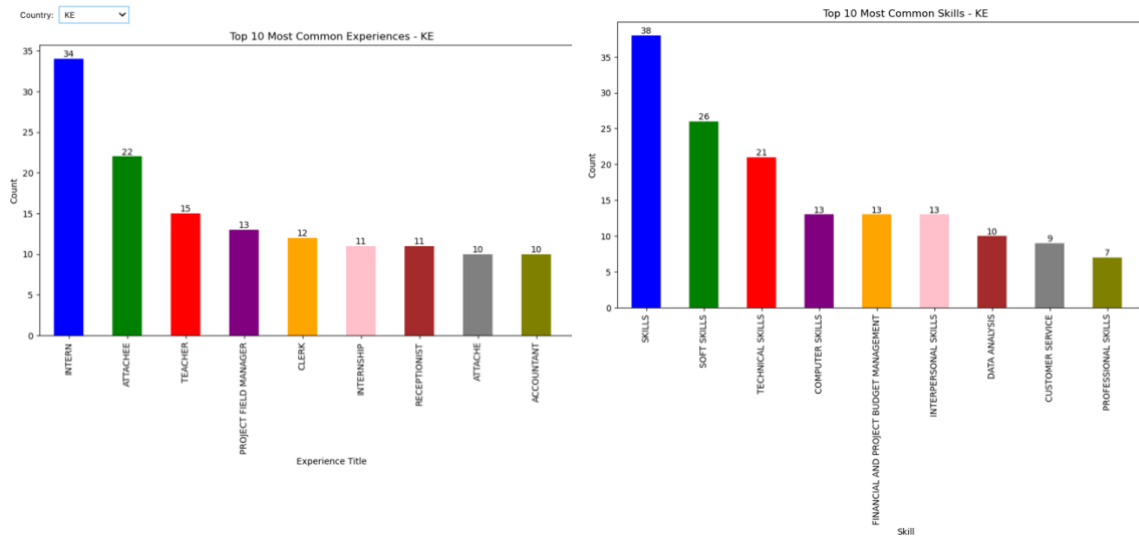
Age Range: 45
Experience Level: BEGINNER
Count: 907

Trend Analysis: Age Range and Experience Level - INTERMEDIATE

Age Range: 45
Experience Level: INTERMEDIATE
Count: 411

Trend Analysis: Age Range and Experience Level - ADVANCED
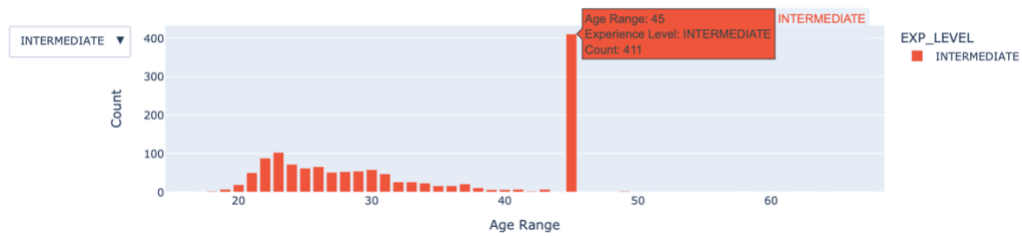
Age Range: 45
Experience Level: ADVANCED
Count: 264

4. Determine trends in experience and skills for these target users.

D. Future Scope: Identifying and suggesting future improvement for good resume building.

2. Milestones:
To measure progress, we have identified the following milestones:

Milestone 1: Data preprocessing completed.
Milestone 2: Dataset ready for analysis.
Milestone 3: EDA conducted, and key insights extracted.
Milestone 4: 3 out of 4 Research questions answered for No job description and With Job description dataset.

3. Job Assignments:
To ensure effective collaboration and task management, the following job assignments have been made:

Abhilash Dikshit: Responsible for data preprocessing, cleaning, conducting EDA and visualizations, Answering Research questions, Identifying future scope.
Kush Patel: Reporting.
Siddharth Alashi: Conducts EDA.

4. Key Risks and Mitigation Strategies:
Identified key risks and corresponding mitigation strategies include:

Risk 1: Insufficient domain knowledge: Collaborate with subject matter experts and conduct thorough research to gain a better understanding of the dataset and domain.

Risk 2: Data quality and missing values: Implement robust data preprocessing techniques and consider appropriate imputation methods to handle missing values effectively.

Risk 3: Model overfitting or underperformance: Employ cross-validation techniques, regularization, and hyperparameter tuning to mitigate overfitting. Evaluate multiple models and select the one with the best performance.

5. Measure of Success:
The measure of success for our project will be determined by achieving the following:

   a) Accurate and reliable predictive or descriptive models that meet the project requirements.
   b) Meaningful insights and actionable recommendations derived from the analysis.
   c) Proof of concept that demonstrates the feasibility and value of the analytic approach.

6. Presentation Method and Delivery of Proof of Concept:
The proof of concept and project findings will be presented through a comprehensive report and an interactive presentation. The report will include an executive summary, methodology, analysis results, key insights, and recommendations. The presentation will utilize visual aids, such as charts, graphs, and interactive dashboards,

III. References:

People analytics. (n.d.). Deloitte Insights. https://www2.deloitte.com/us/en/insights/focus/human-capital-trends/2016/people-analytics-in-hr-analytics-teams.html

Reimagining HR: Insights from people leaders. (2022, March 1). McKinsey & Company. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/reimagining-hr-insights-from-people-leaders