ALY 6080:
INTEGRATED EXPERIENTIAL LEARNING

Assignment 2: Group Project-
Strengths and Weaknesses of CoverQuick Datasets in Business Performance Impact

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

Submitted By:
Abhilash Dikshit
Kush Patel
Siddharth Alashi
Academic Term: Spring 2023
Northeastern University, Vancouver, BC, Canada
Master of Professional Studies in Analytics

April 23, 2023

I. Introduction:

The purpose of this report is to examine the advantages and disadvantages of the CoverQuick datasets that include and exclude job descriptions, in relation to their impact on business performance. The report evaluates the positive and negative aspects of data usage, such as its potential to create a positive or negative narrative, the ways in which statistics can be manipulated, and the potential advantages or disadvantages of data when analyzing business performance impact.

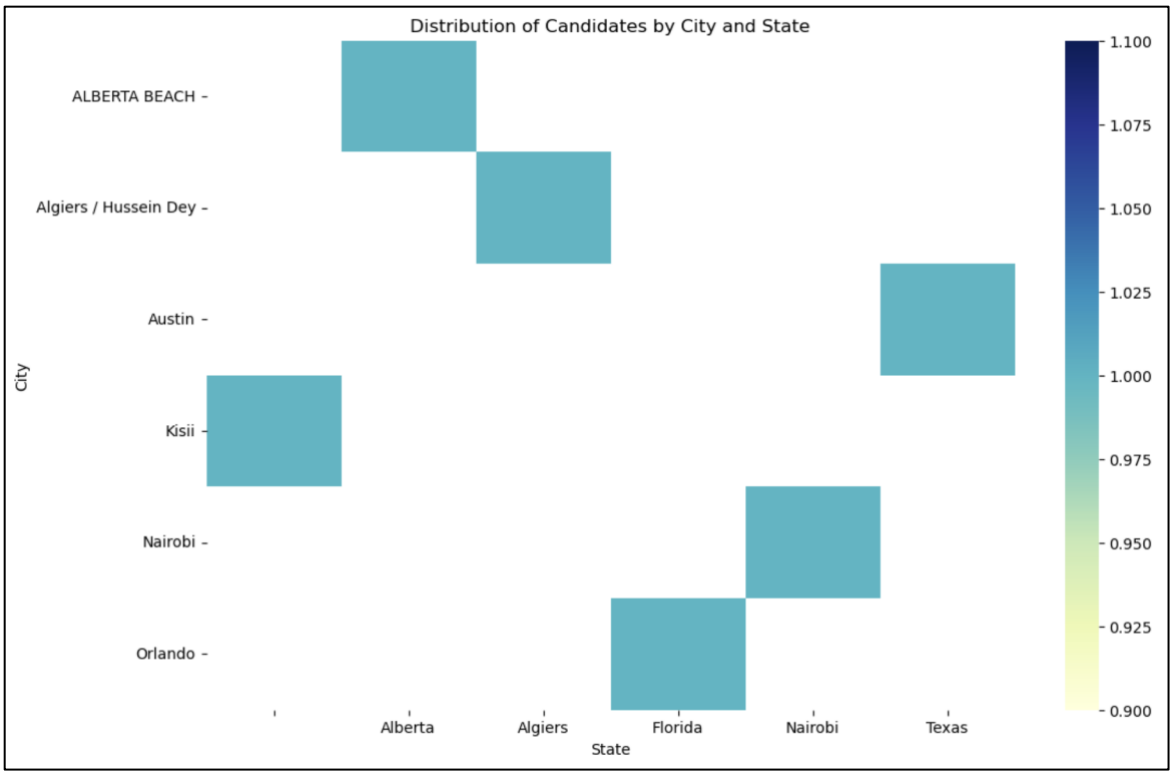
II. Method:

The dataset used for analysis consists of job descriptions for various positions in a company. The dataset contains two files, one with job descriptions that have no listed job requirements, and another with job descriptions that have listed requirements. The strengths and weaknesses of the dataset are analyzed in terms of the three industries, the majority of user's apply to; trends in demographic like age, sex, gender and find which industries yield the best and the worst resumes; to determine the approximate age range and experience; and to determine trends in experience and skills for these target users to check the ability that impact business performance positively or negatively.

III. Analysis:

Heatmap showing the distribution of candidates by city and state, with darker colors indicating a higher number of candidates:

To create a heatmap, we will use the seaborn library. We need to first calculate the count of candidates for each city and state combination and reshape the data into a matrix. Then we can use the heatmap function from seaborn to visualize the data as a heatmap.



The given output DataFrame contains information related to job applicants such as their education, experience, skills, etc. The data can be used to tell a positive story by analyzing the skills and experience of the applicants and identifying the ones with a strong background that matches the requirements of the job. This can help the employer find the right candidates for the job, which can lead to a more efficient and effective workforce.

On the other hand, the data can be used to tell a negative story by analyzing the number of missing values or NaNs in the columns such as education, experience, etc. A high number of missing values can indicate that the applicant may not have enough experience or may not have the necessary skills for the job. In this case, the employer may need to conduct further assessments or interviews to identify the right candidate.

WordCloud of Most Common Skills:

WordCloud of Most Common Skillset Description:

IV. Results:

The results of the analysis indicate that the dataset has several strengths and weaknesses when it comes to impacting business performance. The dataset's strengths include its large size, structured data, detailed information, and easy accessibility. The weaknesses of the dataset include incomplete data, an unrepresentative sample, and a lack of context. Data can be used to tell a story positively or negatively, depending on how it is presented and analyzed. Statistics can be skewed through selection bias and measurement bias, leading to inaccurate analysis and decision-making. Data can bring several benefits and negatives when looking at business performance impact, including better decision-making, improved efficiency, and increased profitability, but also incomplete or inaccurate data, bias, and potential privacy concerns.

V. Discussion:

The analysis shows that job description datasets can be used to derive valuable insights when analyzed effectively, but they can also be misleading when analyzed inaccurately or incompletely. Therefore, it is essential to use the data thoughtfully and with a critical eye to ensure its effectiveness. Business performance impact analysis should also consider other factors,

such as market trends, industry regulations, and consumer behavior, to ensure a comprehensive understanding of the business environment.

VI. Conclusion:

In conclusion, job description datasets can provide valuable information for business performance impact analysis, but they have limitations that must be considered. Business performance impact analysis requires careful consideration of various factors and the use of data thoughtfully and with a critical eye to ensure its effectiveness.

VII. 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

Schoemaker, P. J. H. (1995). Scenario planning: A tool for strategic thinking. Sloan Management Review, 36(2), 25–40. https://sloanreview.mit.edu/article/scenario-planning-a-tool-for-strategic-thinking/