**1. Introduction:**

In 2012, spending per employee remained stable (Miller, 2013). However, with increasing role of technology year by year, organizations struggled with their training policies. Technologies made it easier for employers to improve social learning, but it also increased the costs for employee training. There is also the Solow Paradox theory, which states that in the era of technology, the benefits have not yet been materialized at scale. Despite the high investment in technological innovation, many companies have not seen a direct and immediate impact on productivity growth. This brings into question the role of training on productivity growth, especially when capability gaps is one of the top challenges of corporates, controlling for other factors. The paper hypothesizes that more investment in employee training will increase productivity as it equips the employees with necessary skills and knowledge.

**Data:**

This research looks at the HR Analytics Case Study dataset from Kaggle, which revolves around a fictitious company with a high turnover rate. The data sources contain several datasets, but we will mainly focus on the three datasets called the employee\_survey and draw relevant explanatory variables from the two datasets called general\_data and management\_survey. The merged dataset also requires some data cleaning and removal of variable(s) that we consider are irrelevant, such as whether the employees are over 18 or not, stock option level, standard hours, etc. Our proposed data will contain explanatory variables from the employee\_survey and general\_data, and the dependent variable from the manager\_survey. The general\_data also includes the main chosen variable. The research paper focuses on the employees who are currently still working at the companies, with employee performance score as the indicator for productivity of those employees. Proxies for relevant factors which affect the performance are drawn from those three datasets.

**Variable Selection:**

Using Performance Rating (on a scale of 1 to 4) as an indicator for the response variable productivity and the main chosen variable is the Training Times Last Year employees received, this paper hypothesizes that the more training employees receive, the better they perform at their work, and so the productivity of the company is also improved.

We believe that demographic background (age, education, gender and marital status) can also affect employee performance at work. For example, people who are young and busy with their personal life might do more poorly at work because they lack the experience. It can also be argued that males usually perform better than females because they place a higher emphasis on their careers and have more time for their jobs. People who obtain higher education may perform better than those who have less education because they have knowledge and quick-to-learn ability. However, the correlation may change depending on the education field. Some fields require employees to have more technical skills, and thus can perform better in technical-related jobs than client-oriented jobs. That is why we also include department variable in our dataset.

Regarding individuals’ experience at the company (years at the company, job involvement), people who have worked longer at the company may perform well because they understand the company’s culture. Their level of job involvement might also affect their performance outcome, as the more active they are in the job, the more they learn and excel. Therefore, they might have higher performance ratings from their managers than those who are less proactive or just start to work at the company.

Workplace-related factors (job satisfaction, environment satisfaction) are important because a good working environment motivates employees to perform better at work than those who work under detrimental working conditions, provided that they have the same provision of training.

**2. Literature Review:**

The first paper studies the effect of training on employee performance through effective training programs using qualitative research. Specifically, it reviews the literature and case studies to understand the theoretical framework and models related to the training programs. The researchers conclude that managers do not recognize the importance of training and its effect or employee performance, thus not investing much resources on employee development. However, unless poor performance is caused by the lack of skills, training will have minimal effect on productivity. The researchers also propose a guideline to determine the true causes of the performance problems so that appropriate training programs can bring out the best outcomes. This paper counters the hypothesis that training will always improve productivity by stating that training can only improve employee performance if the problems are caused by insufficient skills and knowledge.

Imran2, A., & Imran, A. (2013). The Effect of Training on Employee Performance . *European*

*Journal of Business and Management* , *5*(4), 137–147. Retrieved from https://pdfs.semanticscholar.org/354c/2c8c60f37f5e25f63f557b3573ec366197ae.pdf

The second paper supports our hypothesis by examining the attitudes of administrative leaders and administrative employees towards the components of training programs and the impact of training on job performance at a university in Jordan. The study includes two questionnaires, the first questionnaire is related to the attitudes towards the components of the training programs, and the second questionnaire is concerned with the relationship between training and performance of employees. It then runs statistical analysis on the data collected to test the hypotheses. The research concludes that there is a positive correlation between impact of training and the performance of employees at the universities. The researcher suggests more suitable training programs for employees and equal opportunities for the employees to attend the training courses.

Al-Mzary, M, et al., (2015). Training and its Impact on the Performance of Employees at

Jordanian Universities from the Perspective of Employees: The Case of Yarmouk

University . *Journal of Education and Practice* , *6*(32). Retrieved from https://files.eric.ed.gov/fulltext/EJ1083504.pdf

The third study also agrees with our hypothesis that investment in employee training can create better employee performance. The researchers explore the impacts of training investment on productivity in 14 Canadian industries from 1999 to 2005. They conclude that in 12 out of 14 industries, training increases productivity. However, they find that productivity does not equal with higher return on investment, as the return on investment was only positive in only four industries. They also suggest that despite the low return on investment, employers still provide employee training because they see the threat of new technology which requires an investment in training to maintain the firm’s current labour productivity. The paper proposes an interesting idea that firms are aware of the impact of training on productivity, especially with the rise of technology, despite the low return on investment. It also challenges our idea of the Solow Paradox that firms do not care about and invest enough in employee training to boost productivity with the rise of technological innovation.

Jennifer, P., Brian, C., & Steven, F. (2013). Return on investment for workplace training: the

Canadian experience, *17*(1). Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/ijtd.12002

**3. Variable Chart:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Name** |  | **Values** | **Units** |
| Employee Survey Data | Employee\_ID | ID Number of the employees | 1 to 4410 | NA |
|  | Environment Satisfaction | Work Environment Satisfaction Level | Low, Medium, High, Very High (1 to 4) |  |
|  | Job Satisfaction | Job Satisfaction Level | Low, Medium, High, Very High (1 to 4) |  |
| General Data  (has 4410 rows) | Age | Age of the employee | 18 to 60 | Years old |
|  | Department | Department in company | Research & Development, Sales, Human Resources | NA |
|  | Education | Education Level | Below College, College, Bachelor, Master, and Doctor | NA |
|  | Education Field | Field of Education | Human Resources, Life Sciences, Marketing, Medical, Other, Technical Degree | NA |
|  | Gender | Gender of employee | Male and Female | NA |
|  | Marital Status | Marital status of the employee | Single, Married, Divorced | NA |
|  | Years At Company | Total number of years spent at the company by the employee | 0 to 40 | Years |
|  | Training Times Last Year | Number of times training was conducted for this employee last year | 0 to 6 | Times |
| Manager Survey | Performance Rating | Performance rating for last year | Low, Good, Excellent, Outstanding (1 to 4) | NA |
|  | Job Involvement | Job Involvement Level | Low, Medium, High, Very High (1 to 4) | NA |

**4. Methodology:**

Our methodologies include conducting literature reviews and performing multilinear regression analysis, looking at the correlation between the main chosen variable Training Times Last Year and the dependent variable Performance Rating while adjusting for other explanatory variables. We plan to do backward model selection to remove the irrelevant variables by examining the t-statistics, F-statistics, and R squares, etc.

After searching on various sources, we decided to choose the datasets from Kaggle since we believe it provides us with the information we need in order to do analysis for our research questions. The link is below

<https://www.kaggle.com/vjchoudhary7/hr-analytics-case-study#manager_survey_data.csv>