

Analyzing and Anticipating the Growth of the Pradhan Mantri Kaushal Vikas Yojana in the Region, Indore using Regression in Python.



Abhishek Varma, Gaurav Kasera

Abstract: *The Indian population crossed 138 Crore numbers and the second-most populous nation in the world. Recent changes in job structure show that Indian labor is transitioning from the farming sector to the industry-and-services industries. India is unique globally, thanks to the demographic dividend compared to developed countries with the aging population. India's 65 percent workforce is under the age of 35, and by 2025 70 percent of the country would be working-age. A large population, if educated, qualified, and productive, will effectively escalate the value of the demographic dividend and leading to sustainable development. Prime Minister Narendra Modi has unveiled a training program for India on July 2015 and program aims to train about forty lakh trainees in India with various cognoscendi by 2022. This is a revolutionary program aimed at offering realistic, industry-germane, and skills-based training to young people for the last four years. It is crucial to examine the role and influence of Pradhan Mantri Kaushal Vikas Yojana (PMKVY) at the outset and how it will help people coalesce these skills and become employed. The data of the Indore region obtained from the official site is analyzed with the help of python and shows whether PMKVY buttresses the Indian economy or not through tests performed. By linear regression future expectancy of the program is predicted.*

Keywords: *Skill development, skill India, python, Linear regression, Pradhan Mantri Kaushal Vikas Yojana.*

I. INTRODUCTION

Pradhan Mantri Kaushal Vikas Yojana (PMKVY) is the Government of India's skill growth program system for capability identification and standardization. The cornerstone of the PMKVY scheme is to simulate skills towards employable skills and to ameliorate the job productivity of the country. Trainees were benefited by giving them cash and also spurred with a credential certificate valid for two years following completion of the course. India has an opportunity to become a global provider of skills with a surfeit of skilled labor, especially the aging of developed nations. Earlier few Indian workers had any coordinated skills training. Nor is it surprising that there is a dearth of skilled workers in particular areas of the country's economy. At the same time, profuse of the country's youth is searching to face economic and livelihood subsistence.

Revised Manuscript Received on September 13, 2020.

* Correspondence Author

Abhishek Varma*, Department of Electrical Engineering Shri Vaishnav Vidhyapeeth Vishwavidyalaya Indore, Madhya Pradesh, India. Email id:-abhishekvarma827@gmail.com

Gaurav Kasera, Department of Civil Engineering Shri Vaishnav Vidhyapeeth Vishwavidyalaya Indore, Madhya Pradesh, India. Email id:-gakasera4@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

It is necessary not only for economic extension but also for aid. Fulfill the expectations of young people for high education, higher-paying jobs, and self-employment prospects. Moreover, it will also encourage the nation to take advantage of its favorable demographic profile.

The total value of the award per individual was ₹8,000 (US\$ 110). The government will recognize those wage earners who already have a standard level of skill as per the scheme. A target to distribute approximately ₹115 billion (US\$ 210 million) is set for the scheme in the first year. Education programs were drawn up based on National Occupational Standards (NOS) and training packets established explicitly in various expertise sectors. Various Business Competence Councils (SSC) produced with the involvement of Businesses have established training plans and quality plans for this. The National Council for Skill Development (NSDC) has made a coordinating and driving agency for the same purpose, would improve the country's competitiveness, and Provide workers with high-quality instruction in expertise across a wide variety of sectors.

Instead of antiquated tools, PMKVY comprises high tech tools and labs, which is imperative to enhance skills. Furthermore, make trainees explicit on the subject. Until python language, there is a lot of problem managing massive data, but it has rendered the study, processing, and simulation quite simple, and more python has a range of excellent libraries from which users can build highly personalized plots. A bit of an overview of the few popular libraries we used, Matplotlib: low level-offers much versatility, Pandas Visualization: simple to use GUI (Graphical User Interface), based on Matplotlib. It comes with all the tools needed for data visualization offerings.

II. LITERATURE REVIEW

“India is the second most populated country in the world with nearly 135 crore population according to the 2019 revision. Out of this more than 65% of the population falls under the age of 35 years, which is the main workforce of any country. It is a major task for the country to implement, administer and monitor the scheme to skill this section of the population for an effective and positive outcome. Jan Shikshan Sansthan (JSS) is conceived as an establishment for conducting skill up-gradation vocational training in the areas of Non-formal, Adult and Continuing Education programme in both industrial/urban and rural areas. Main objectives of JSS's are to impart vocational skills, life skills and technical knowledge to neo-literates to raise their efficiency and productive ability to supplement their income level.



This paper presents the features of JSS Portal and how technology supports and enhances the implementation of Scheme of Jan Shikshan Sansthan.”[1]

“The productivity of Indian labour is lower as compare to the Asian economies and need up-skilling of the labour to ensure economic growth.”[2]

“Identified two main ways for developing skills that are through education & training and developmental experiences which is significant for the economic development of the country. Skill is the capability to do something good to arise from talent and training. Also Author state the skill development model and training methods for micro-entrepreneurs in India, Model fall into three categories that are general management, entrepreneurial and knowledge related skill. Unemployment problem is the major challenge to India, youth enter in job market every year after completing their education but remain jobless due to lack of opportunities. Promoting self-employment entrepreneurial among unemployed youth is a good solution. With this constraint, the model was developed to provide training to unemployed youth to engage in wages employment to earn a livelihood .”[3]

“To contemplate mindfulness about expertise India battle among youth, To consider the response of crusade from youth, to examine major and minor factor influencing the battle, to ponder the confinements and container necks which are coming in the crusade, to think about the effect of battle on work generation .”[4]

“Clarified that training is a significant instrument to improve individual efficiency and eventually will boost the value of organizations but organizations should strike the equilibrium between education and instruction. The final findings showed that the impact of training in various industries is varying.”[5]

“Appropriate from the earliest starting point of freedom, the Indian policymakers have expected to advance reasonable monetary improvement in the nation. This includes adjusting financial development which includes adjusting monetary development with the diminished quickening of unevenness and expanded social insurance of destitute individuals. Welfare conspires in India and their impact on social affirmation amid a period of high financial development instructs us concerning India's execution in key social welfare zones. India is spending less on social welfare programs and other welfare designs than various countries, centre salary nations monetary execution has surpassed its execution in social welfare pointers. The manageability of the welfare program will increment if viability and productivity are expanded.”[6]

“Regression analysis should be viewed as a set of data analytic techniques used to study the complex interrelationships that may exist among variables in given environment.” [7] “Python encourage clean code design ,high level structure, and “packaging” and faster development time required as projects expand in breadth and scope.Where you can grow your code from project to project, add other new or existing python elements, and reuse code at any whim.” [8]

III. RESEARCH METHOD

This research results in the PMKVY trainee group trained at the PMKVY training center under India's cleanest region, Indore. In this, we have analyzed the impact of PMKVY on Indore, and due to the present condition of 2020 pandemic (COVID19), instead of taking primary data, we have relied on secondary data from numerous Publications, Books, Papers, Documents, Records of State, Theses of different thinkers, Etc. The logical meaning is developed and based on information gathered from different secondary resources. Selecting the appropriate data set is a daunting job, and this may exclude data due to which data may not be homogeneous, and this affects the overall accuracy of the result to deal with data collected from the official website of having 37402 students enrolled in 31 different sector from December 2016 to till February 2020 and the data was last updated on 30 June 2020. It shows that there was no neophytes student enrolled after February 2020. We were using **plt. Scatter** library and assigning the input variables we got this plot. Here we have obtained sporadic output as the number of students enrolled and placed differs over the year. The X-axis shows the total working duration(in months) of the program, and Y shows the number of trainees enrolled and placed over the month.

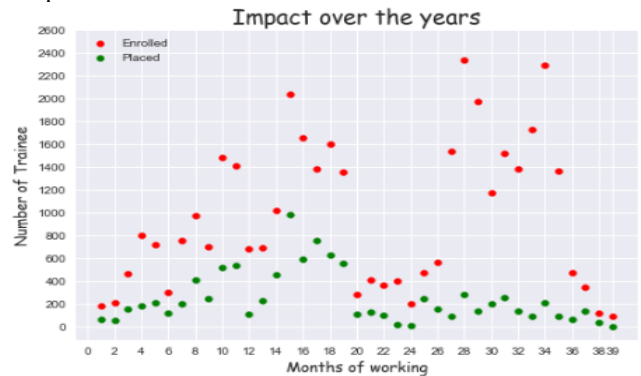


Figure: 1 Shows the scatter plot having a linear relationship between the variables.

Table: 1 Demographic Characteristics of PMKVY Trainee

Demographic Characteristics	Classification	Male	Female	Frequency	Place
Sector	Agriculture	557	1382	1939	1045
	Apparel	3211	154	3365	1460
	Beauty and wellness	3793	217	4010	1248
	Electronics and hardware	857	1667	2524	1200
	IT-ITes	1271	658	1929	475
	Retail	3899	8492	12391	1046
	Telecom	1336	1480	2816	1015

A. Tools used in data calibration:

Python has become a game-changer in Data science. It is a nifty tool with several advantages. It is flexible and continually improving because it is open-source, and it has a plethora of useful libraries by this one can visualize, Several data tests are analyzed and performed.



B. Hypothesis:

- H0** - Number of enrolled students $\leq 20\%$ increase in employment over the year. Moreover, H0 rejected when the chi-square value exceeds in comparison with the chi-square distribution table and vice-versa.
- H1** - Number of enrolled students $\neq 20\%$ increase in employment over the year. Moreover, H1 is Accepted when the chi-square value exceeds in comparison with the chi-square distribution table and vice-versa.

C. Factor analysis:

The P-value obtained from the chi-square test for independence is significant ($P < 0.01$), and therefore, we infer that there is a noteworthy association between the two input variables. To get chi-square value from the data we have used-

`chi2.p_value, dof, exp=stats.chi2_contingency(categories)`

where categories= data

chi-square value, p-value, degree of freedom values are obtained.

We obtained a Chi-square value of 569.837, and P-value is less than 0.001. Here we have taken a 99.99% confidence level, and the significance level is 0.001. The degree of freedom is 38 and level of significance (α) is obtained as $\alpha < .001$

For every single possibility, the chi-square table offers no exact values. We will get an exact number if we use a calculator. On comparing Calculated Chi-Square value with 73.402, which is at 38(dof) and 0.001(α), and based on the result, we **rejected the Null hypothesis**.

Table: 2 Factor analysis

Calculation through python using scipy.stats library.		Sample Size 37402
H0 ≤ 73.402 (Rejected) H1 $\neq 73.402$ (Accepted)	Chi square test	569.837
	P value	>0.0001
	Degree of freedom	38
	Level of significance	0.001
	Level of confidence	0.999

D. Data analysis:

Linear regression is a statistical model that implies a linear relationship between the dependent variable (which is the predictor that we are attempting to predict / estimate) and the independent variable(s) used in the analysis. Under Simple Linear Regression, one can predict the dependent variable by utilizing just one independent/reference variable. It has the framework of the following:

$$Y = C + M * X$$

Y = variable depending on (output / outcome / projection / estimate)

C = Constant (Y-Intercept)

M = Regression line slope (the effect X exerts on Y)

X = Independent component (input component used in Y forecast)

Regression is searching for relations among the variables. The dependent feature here is "placed" and also regarded as dependent variables, outputs placed on the 'Y' axis and the independent feature here are 'months' placed on the 'x' axis and is also regarded as independent variables or inputs. However, the inputs of the data can be discrete, continuous, or even categorical data such as gender, sectors, Etc. The denotation of outputs with y and inputs with x is standard procedure. We want to predict Y's value using regression analysis, provided that we have the value of X. There is a causal relationship between the two. It shows how much y changes to x for each function. The more the student enrolled, the higher the employment rate we will likely receive. Researchers dropped unwanted columns to make data simple for regression analysis, and columns having a relationship with each other is considered, which will cause simplification while running codes. We did the Formatting of data as per input given, and then it is sorted. The regression line is the line that better matches the data points. The 'green' points that are scattered are the observed values. M is the Regression line slope.

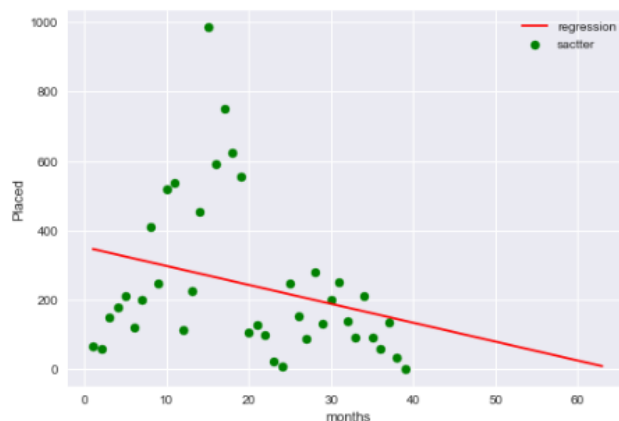


Figure: 2. Data analysis(Regression analysis)

```

1 mean_x = np.mean(X)
2 mean_y = np.mean(Y)
3 n = len(X)
4 numer = 0
5 denom = 0
6
7 for i in range(n):
8     numer += (X[i]-mean_x) * (Y[i]-mean_y)
9     denom += (X[i]-mean_x)**2
10 b1 = numer/denom
11 b0 = mean_y - (b1 * mean_x)
12
13 max_x = np.max(X)+24 #Calculated 24 by putting values when x=6months of working then y=0
14 min_x = np.min(X)
15
16 #Calculating line values x and y
17 x = np.linspace(min_x,max_x,1000)
18 y=b0 + b1*x
19
20 #Plotting line
21 plt.plot(x,y, color='r',label='regression')
22
23 #Plotting scatter plots
24 plt.scatter(X,Y,color='g',label='scatter')
25 plt.xlabel('months')
26 plt.ylabel('Placed')
27 plt.legend()
28 plt.show()
29
30 #future placed = -5.4449*months + 351.668
    
```

Figure: 3 Coding of regression analysis

E. Performance of the model:

We have to split data into training and testing and then take 80% of all the data and then taken to train the model. After model fitting, the model result is viewed by the model to predict actual(True) data by comparing model prediction.



Analyzing and anticipating the growth of the Pradhan Mantri Kaushal Vikas Yojana in the region, Indore using regression in python.

An important thing to remember is that the training should not see the model as the aggregate data during training. R-square offers the goodness of fit.

It represents a measure of how the model will possibly estimate test samples by the proportion of the overall r-square value described in the variance is one, and this is also known as the determinative coefficient. The existence of outliers will burst the term of the error to a very high value. Nevertheless, in the case of RMLSE(Root means square log error), the outliers are dramatically scaled-down, and RMLSE expunges their effect.

IV. PROPHECY

The government shut down the skill development programs due to the outbreak COVID-19 pandemic in the year 2020, but according to our theory Of regression, the total working months from the start are 64 months from December 2016 to March 2022. If the government continues the program at the same pace as before, then it will be in existence until March 2022. we have calculated that the total working months from the starting is of 64 months, and it was from December 2016 to March 2022.

Table: 3 OLS Regression
OLS Regression Results

Dep. Variable:	R-square : 0.076
placed	
Model : OLS	Method: Least square
A I C : 533.2	Prob (F - statistic) : 0.0897
B I C : 536.5	Log - Likelihood: -264.60
Df Residuals : 37	Df Model: 1

Intercept(c) = 351.668

Slope(m) = -5.444

$Y = m(x)+c$

$Y = -5.444(64)+351.668$

$Y = 3.449$ is the closest value to 0.

V. RESULT

We carried out this research to find the future hope and job levels over the year of the system in Indore, and according to our analysis, the program will be maintained by the government until March 2022. However, due to pandemic, this will only last until February 2020, which is the intricate problem. That is a preeminent concern because this initiative is mostly pragmatic as there is no other such initiative yet to enhance & cultivate the skills and abilities of youth. This plan has finished two years before causing an abysmal situation; this will do serious harm to young people who cannot profit from this scheme. As per the demographic profile and low literacy rate of the largely uncertified workforce, this is an insurmountable problem to jobs in the global manufacturing sector, so the government has to strive at other programs to assuage and constraint the adverse effects.

VI. CONCLUSION

The youth of India is facing many challenges, including

Unemployment, poverty. As a result, India's government is taking some measures to reduce these issues by introducing a few steps to build abilities and self-reliance. Developing India is taking a big step by inspiring young people's potential to make the future bright. Such numerous programs allow young people to boost the standard of their skills and re-skilling in various fields through a training system that immediately improves the growth of personality and overall progress of young people. People have been endeavoring for the job, and the boon situation is not more there.

VII. LIMITATION OF THE STUDY

1. Our analysis is restricted to the chosen Indore city area. Consequently, the analysis findings can not be generalized.
2. There are drawbacks to the statistical tools used to interpret the results.
3. Many of the drawbacks apply to this analysis of primary results.

REFERENCE

1. S. Kanchan and S. Varshney, "Skill development initiatives and strategies", *Callcentre.ddugky.info*, 2015. [Online]. Available: http://callcentre.ddugky.info/dms/documents/1529234017_EIJMRS5242.pdf. [Accessed: 17- Jul- 2020].
2. L. Sharma and A. Nagendra, "Skill Development in India: Challenges and Opportunities", *Indian Journal of Science and Technology*, vol. 9, no. 48, p. The productivity of Indian labour is lower as compare to the Asian economies and need upskilling of the labour to ensure economic growth, 2016. Available: 10.17485/ijst/2016/v9i48/107324.
3. H. Katole, "SKILL DEVELOPMENT AND ECONOMIC GROWTH OF INDIA", *Aebjournal.org*, 2020. [Online]. Available: <http://www.aebjournal.org/articles/0304/030401.pdf>. [Accessed: 17-Jul- 2020].
4. R. Dangwal, K. Sharma and S. Hazarika, "Hole-in-the-Wall learning stations and academic performance among rural children in India", *Journal for Multicultural Education*, vol. 8, no. 1, pp. 31-53, 2014. Available: <https://www.emerald.com/insight/content/doi/10.1108/JME-03-2013-0006/full/html>.
5. S. Javed, S. . and M. ., "Impact of Training on Expectation of Employee and Employer: A comparative study", *Information Management and Business Review*, vol. 5, no. 12, pp. 601-607, 2013. Available: 10.22610/imbr.v5i12.1094.
6. R. Jha, "Welfare Schemes and Social Protection in India", *SSRN Electronic Journal*, vol. 3434, 214-231, no. 0144-333, p. Available: 10.2139/ssrn.2303988.
7. S. Chatterjee and A. Hadi, *Sensitivity analysis in linear regression*. New York: Wiley, 1988, p. 6.
8. W. Chun, *Core PYTHON programming*. Harlow: Prentice Hall, 2006, p. 5.

AUTHOR PROFILE



Abhishek Varma, obtained his diploma (Electrical) from Shri Vaishnav polytechnic college, Indore and he is currently pursuing his bachelor of technology(Electrical Engineering) From Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore meanwhile he has completed several internships from reputed institutions like NSDC, MPPGCL. Speed measurement using an IR sensor, line follower robot, function generator were his projects.



Gaurav Kasera, has completed his diploma in civil engineering from Shri Vaishnav Polytechnic College and worked on the project on analytical design for the elevated water tank with filter media and pursuing his graduation from Shri Vaishnav Vidyapeeth Vishwavidyalaya.

