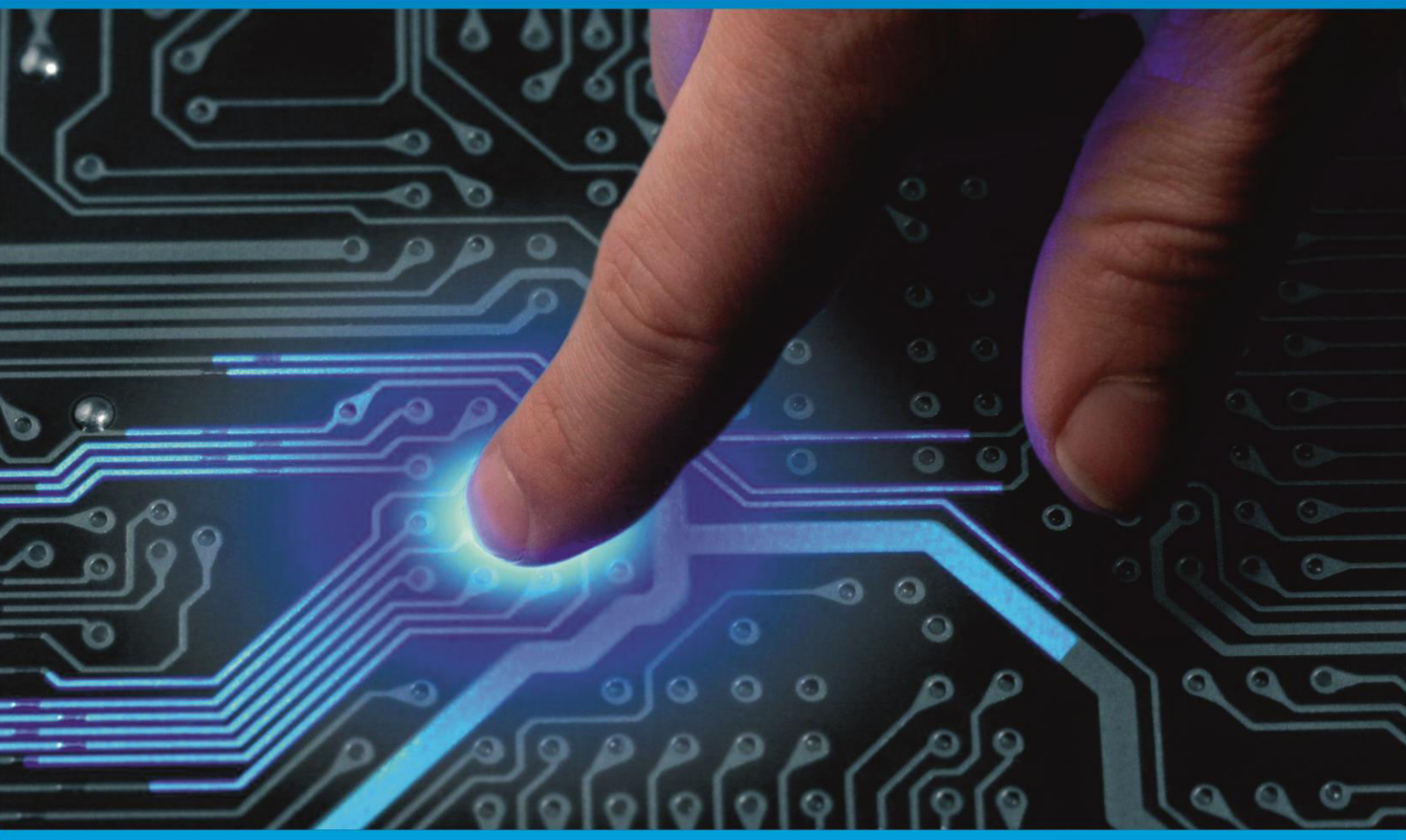
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

A Novel Study and Analysis on Diabetes Patients Vs Non-Diabetes Patients

Abhishek Yernagula[1], Siddharth Nanda[2]

U.G Student, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India [1]

Faculty, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India [2]

ABSTRACT: Diabetes is a chronic health problem with devastating, yet preventable consequences. It is characterized by high blood glucose levels resulting from defects in insulin production, insulin action, or both.1,2 Globally, rates of type 2 diabetes were 15.1 million in 2000,3 the number of people with diabetes worldwide is projected to increase to 36.6 million by 2030.4 In 2007, 23.6 million people, or 7.8% of the United States population had type 2 diabetes. Of these, 90-95% of these cases were adults with type 2 diabetes. Type 2 diabetes impacts men and women proportionately; there are over 12 million men with diabetes and 11.5 women with diabetes. In adult patients, 6.6% were non Hispanic White, 11.8% were non Hispanic Black, 10.4% were Hispanic, and 7.5% were Asian. This rate is expected to increase greatly over the next halfcentury. Along with the increase in incidence of diabetes, both individual and societal expectations concerning the management of diabetes have also increased, with many reports from The Centres for Disease Control (CDC), United States Department of Health and Human Services (USDHHS), and the National Institutes of Health (NIH) urging patients to "Take Charge of Your Diabetes"5 and "Conquer Diabetes".6 One of the main goals of USDHHS's report, Healthy People 2010, is to improve the quality of life for persons with diabetes.7 Taking control of diabetes to improve quality of life has put the spotlight on the need for additional support and education for patients with type 2 diabetes

KEYWORDS: diabetes, Hispanic, insulin.

I. INTRODUCTION

Diabetes mellitus (DM) is most devastating, chronic, common non-communicable disease (NCD) and has become a serious problem globally. The number of the diabetic population around the globe is continuously increasing with a current estimation of 371 million cases in 2012 and it is expected to reach 552 million by 2030 (IDF diabetes atlas, 2012). It is also estimated that 5% of all deaths in the world are caused by diabetes and the number is rapidly increasing. Among two general types of diabetes, type 1 diabetes (TIDM) is characterized by immune complex mediated attack on insulin producing β cells of the pancreas (Atkinson and Eisenbarth, 2001). Type 2 diabetes (T2DM) arises due to either insufficient insulin synthesis or the body's inability to respond secreted insulin and leading to glucose build-up in the blood (DeFronzo, 1997). The impairment in glucose control leads to both micro and macro vascular complications that often result in the other clinical conditions associated with diabetes.

We have taken a dataset from kaggle.com consisting of various attributes such as insulin,age,plasma,etc. of patients tested for diabetes and different hormones in the body of the patients with their age were categorized into two parts i.e.patients who were tested positive in the test vs patients who were tested negative in the test. We have used Systematic Sampling to create some stratas within the dataset to have a uniform data. After creating those strata, we applied some statistics to it, to find the average value of math score needed for different levels of education as well as vaiance of every strata.

Random samples are then selected from each of these stratums. Stratified sampling provides better inclusion of the population. Using these methods we will find out the performance of different students categorized by gender and degree. As my dataset can be divided into groups, I used stratified sampling so that every group can participate equally into the result.
Then, the visualizations are done in Data Studio to show data in graphs,maps,etc. .

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

II. LITERATURE SURVEY

In this chapter, the scientific literature will be reviewed to increase the understanding of the health problem studied, the theoretical framework for the research questions and the empirical evidence for the study framework. The first section of this chapter provides an overview of the epidemiology of diabetes. In this section, the prevalence of diabetes and its complications will be reviewed. The second section provides an explanation of and the relationship between the key theoretical constructs in the conceptual framework of the study. In section three, the major theoretical approaches to this problem are discussed and several behavioral theories and models are reviewed, and the strengths and limitations of these theories and models are addressed. The last section of the chapter synthesizes the empirical evidence and theories to develop a framework for the study. The rationale for the use of such theories is presented as well.

1.**The Incidence of Diabetes**

By: Charles Gbollie and Harriett Pearl Keamu, accessed on 25/03/2021.

Historically, achievements in public health over the 20th century have shifted the focus from communicable diseases to chronic diseases such as diabetes. Globally, rates of type 2 diabetes were 151 million in 2000.1 In 23 North America, 14.2 million people had type 2 diabetes in the year 2000, and that number is expected to increase 23% to 17.5 million people by the year 2010.1 In 2005, diabetes affected 7% of the United States population and was the sixth leading cause of mortality.2 This rate is expected to increase tremendously over the next half-centuryStudent Academic Performance: The Role of Motivation, Strategies, and Perceived Factors Hindering Liberian Junior and Senior High School Students Learning.

2.**Explanation of and the relationships between key study constructs**

By: Rosemary WahuMbogo, accessed on 25/03/2021.

The following section explains the constructs of illness identity, social identity, goal setting, support group participation and goal achievement and the relationships between them. Explanation of these constructs is conducted through selected published definitions and by demonstrating the applications of each within research studies that have been conducted in the literature. This section describes the constructs from a broad perspective that is not limited to the application of such constructs in the health care perspective. Antecedent Factors Affecting Academic Performance of Graduate Students at the Nairobi Evangelical Graduate School of Theology.

3.**Illness identity**

By: ErumShahzadi, Z. Ahmad, accessed on 25/03/2021.

Chronic illness is a “state of unwellness produced by disability or disease requiring medicosocial intervention over an extended interval and affecting many aspects of an individual’s life.”22 A meta-analysis of the literature revealed that the themes of research on chronic illness have shifted from suffering, loss, biographic disruption and sick role to being courageous, maintaining hope, restructuring illness, reframing and reshaping the self, regaining control, redefining health, empowerment, transformation, and normality a decade later. There has also been a focus on the expertise of chronically ill patients in obtaining information about their illness and managing the illness and competence in health care decision-making.22 Chronic illness can influence 28 social relationships-strengthening the relationships between family members in the care-giving process, creating new relationships with other chronically ill patients, and also changing the intensity of the relationship between the patient and the health care provider. It is important to understand the implications of chronic illness on identity.

4.A Role of Glucose Overload in Diabetic Cardiomyopathy in Nonhuman Primates

By: YongjunShen, accessed on 25/03/2021.

Given the fact that T2D is a major risk factor for HF and hyperglycaemia is strongly associated with diastolic dysfunction in diabetes patients, summarized in Figure 1, and the investment in T2D and HF research by the whole society, reliable preclinical models of human disease are urgently needed. NHPs are such an invaluable translational model in the study of human DCM pathology. First, NHPs show the greatest similarities to the disease of humans. Second, the genomes of the commonly used NHPs in biomedical research have been sequenced. Third, NHP models served as unique models for establishing the safety and efficacy of novel drug development in humans.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

5.Stratified Sampling.
By: Van L. Parsons, accessed on 25/03/2021.
This paper tells us about stratified sampling, its uses, its advantages, disadvantages and how it is better than simple and cluster sampling. Also this technique is used to increase efficiency of a sample as well as it is used in the modern technology a lot.

6.Study on a Stratified Sampling Investigation Method for Resident Travel and the Sampling Rate.
By: YongjunShen, accessed on 25/03/2021.
This paper includes some surveys done in two cities using stratified random sampling and it tells the characteristics of travelling in both the cities.

7.Diabetes mellitus: The epidemic of the century
By: Roy Levin, YaronKanza, accessed on 25/03/2021.
Diabetes mellitus is the epidemic of the century and without effective diagnostic methods at an early stage, diabetes will continue to rise. This review focuses on the types of diabetes and the effective diagnostic methods and criteria to be used for diagnosis of diabetes and prediabetes. Evidently, diabetes is a complex disease with a large pool of genes that are involved in its development. The precise identification of the genetic bases of diabetes potentially provides an essential tool to improve diagnoses, therapy (more towards individualized patient targeted therapy) and better effective genetic counseling. Furthermore, our advanced knowledge of the association between medical genetics and the chronic complications of diabetes, will provide an additional advantage to delay or eradicate these complications that impose an immense pressure on patient's quality of life and the significantly rising cost of health-care services.Stratified Sampling of Neighborhood Sections for Population Estimation. By: Roger Hillson, Joel D. Alejandre, Kathryn H. Jacobsen, Rashid Ansumana, Alfred S. Bockarie, UmaruBangura, Joseph M. Lamin, and David A. Stenger, accessed on 25/03/2021.

8.Type 2 Diabetes Mellitus: A Review of Current Trends
By:Abdulfatai B. Olokoba, Olusegun A. Obateru, Lateefat B. Olokoba
Type 2 DM is a metabolic disease that can be prevented through lifestyle modification, diet control, and control of overweight and obesity. Education of the populace is still key to the control of this emerging epidemic. Novel drugs are being developed, yet no cure is available in sight for the disease, despite new insight into the pathophysiology of the disease. Management should be tailored to improve the quality of life of individuals with type 2 DM.

9.IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040
540 data sources were reviewed, of which 196 sources from 111 countries were selected. In 2015 it was estimated that there were 415 million (uncertainty interval: 340–536 million) people with diabetes aged 20–79 years, 5.0 million deaths attributable to diabetes, and the total global health expenditure due to diabetes was estimated at 673 billion US dollars. Three quarters (75%) of those with diabetes were living in low- and middle-income countries. The number of people with diabetes aged 20–79 years was predicted to rise to 642 million (uncertainty interval: 521–829 million) by 2040.
Diabetes prevalence, deaths attributable to diabetes, and health expenditure due to diabetes continue to rise across the globe with important social, financial and health system implications.

10.Prevalence of diabetes, management and outcomes among Covid-19 adult patients admitted in a specialized tertiary hospital in Riyadh, Saudi Arabia
T2DM and hypertension were the most common comorbidities in both males and females, accounting for 45.7% and 28% of all patients included in the dataset, respectively. These figures are much higher at the national level (7.6% and 8.8%, respectively) [6] and this is expected since the present data set deals only with hospitalized patients with moderate to severe symptoms of Covid-19. Both comorbidities however were not associated with increased likelihood of death. Nonetheless, the high prevalence of T2DM and hypertension among hospitalized Covid-19 patients have already been observed not only in majority of industrialized countries with high cases of SARS-CoV2 infection, but have also been associated with higher risk for severe outcomes, including death
Another highlight in the present study is the high mortality rate (10%) of the hospital relative to the national mortality rate in Saudi Arabia which is around 1.5% at the time of this writing. This is expected since only

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

moderate to severe Covid-19 cases are admitted in the hospital and those with no to mild symptoms were advised to self-isolate. Separating males alone and stratifying according to ethnicity, Saudi men had the highest mortality rate compared to South Asians and others, but not statistically significant even before adjusting for age.

III. PROPOSED ANALYSIS APPROACH

Using sample data, estimate the mean math score in the population. Find the margin of error and the confidence interval. Assume 95% confidence interval.

1. **Population Parameter:**

To compute the overall sample mean, we need to compute the sample means for each stratum.

$\bar{x} = \sum(x_i)/n$

Mean (Pregnant) = $\bar{x}_{pregnant} = \frac{335}{71} = 4.71$

Mean (Plasma) = $\bar{x}_{plasma} = \frac{9894}{71} = 139.35$

Mean (Age) = $\bar{x}_{age} = \frac{2490}{71} = 35.07$

After calculating mean of every stratum, the overall sample mean is:

Sample Mean = (4.71+139.35+35.07) / 3

$= \frac{179.13}{3} = 59.71$

Thus, based on the data from the sample strata, we estimate that the mean diabetes(+ve) in the sample is equal to 59.71.
To compute the overall sample mean, we need to compute the sample means for each stratum.

For negative;
$\bar{x} = \sum(x_i)/n$

Mean (Pregnant) = $\bar{x}_{pregnant} = \frac{267}{71} = 3.76$

Mean (Plasma) = $\bar{x}_{plasma} = \frac{7697}{71} = 108.40$

Mean (Age) = $\bar{x}_{age} = \frac{2288}{71} = 32.22$

After calculating mean of every stratum, the overall sample mean is:

Sample Mean = (3.76+108.40+32.22) / 3

$= \frac{144.38}{3} = 35.22$

Thus, based on the data from the sample strata, we estimate that the mean diabetes(-ve) in the sample is equal to 35.22.

Population Mean = $(N_h/N)*x_h$

$= \left(\frac{222}{1000} * 64.32\right) + \left(\frac{118}{1000} * 70.41\right) + \left(\frac{59}{1000} * 70.2\right)$

$= 79.79$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

Based on the data, we can estimate the mean math score for the population is equal to 79.79.

2. Population Variance

We need to compute sample variance within each stratum, to compute standard error for every stratum.

$s^2_{pregnant} = 0.00014$

$s^2_{plasma} = 0.0676$

$s^2_{age} = 0.000625$

3. Standard Error

The standard error measures the variability of our sample estimate of the population mean. We will use standard error to compute the margin of error and to define a confidence level.

$SE = (1 / N) * sqrt \{\Sigma [N^2_h * (1 - n_h/N_h) * s^2_h / n_h] \}$

For +ve,
$SE_{pregnant} = 3.83$
$SE_{plasma} = 30.94$
$SE_{age} = 9.22$

Thus, the standard error of the sampling distribution of the mean is 14.66.

For -ve,
$SE_{pregnant} = 3.15$
$SE_{plasma} = 30.72$
$SE_{age} = 11.78$

Thus, the standard error of the sampling distribution of the mean is 15.21.

We are working on 95% confidence level.

4. Confidence Level

As the sample size is large, thus we will use z-test. Standard Normal Distribution Table is used to find critical z-score. In this part of the analysis, usually researchers choose a confidence level and the most frequently chosen confidence level is 95%. Thus, we will use that only.

5. Critical Value

The critical value is a factor used to compute the margin of error. To find the critical value, we take these steps:

- Alpha(α):
  α = 1-(confidence level/100)
  α = 1-(95/100)
  = 0.05
- Critical Probability (p^*) :
  p^*= 1- (α/2)
  = 1-(0.05/2)
  = 0.975

Using Standard Normal Distribution Table, we can see that the critical z-score value is 1.96.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

6. Margin of Error

For +ve;
ME = (Critical Value * Standard Error)
= 1.96 * 1.738
= 3.40
For -ve;
ME = (Critical Value * Standard Error)
= 1.96 * 1.78
= 3.50

7. Confidence Interval

The minimum and the maximum values of the confidence interval are:

CI_{min} = x – Standard Error * Critical Value
CI_{max} = x + Standard Error * Critical Value

For +ve;
CI_{min} = 59.71– (14.66* 1.96)
= 30.98
CI_{max} = 59.71+ (14.66* 1.96)
= 88.44
For -ve;
CI_{min} = 35.22– (15.21* 1.96)
= 5.41
CI_{max} = 35.22+ (15.21* 1.96)
= 65.03

Summary

Based on our sample data, we estimate that our population mean is 79.79. given a 95% confidence level, the margin of error around that estimate is 3.5; and the 95% confidence interval is from 30.98 to 88.44. Thus, we can see that the population mean is between the confidence interval which means an average of 77.79 math score is required for different levels of diabetic patients.
The data is shown in Data Studio through bar graphs and scorecards,etc. .

Visualization

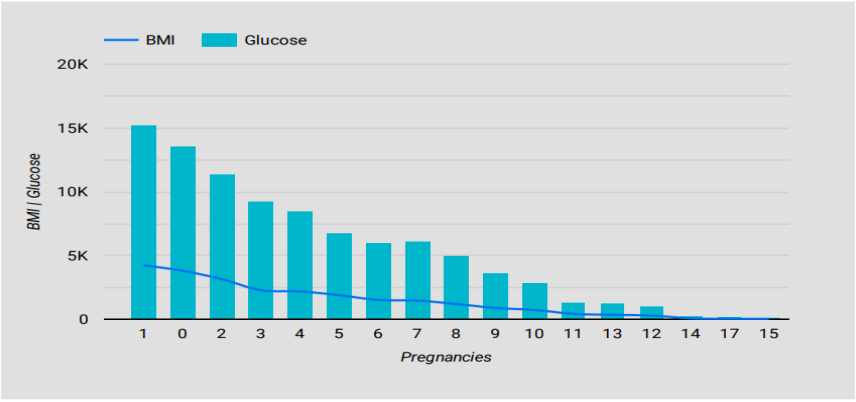


Fig 1: Bar Graph for Glucose vs Pregnancies

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

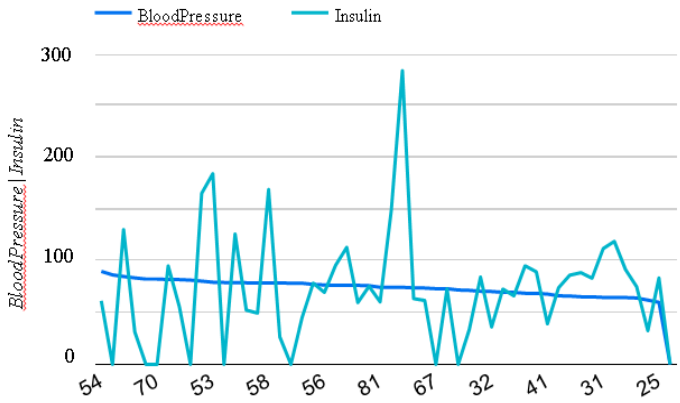


Fig 2: Graph Insulin vs Age

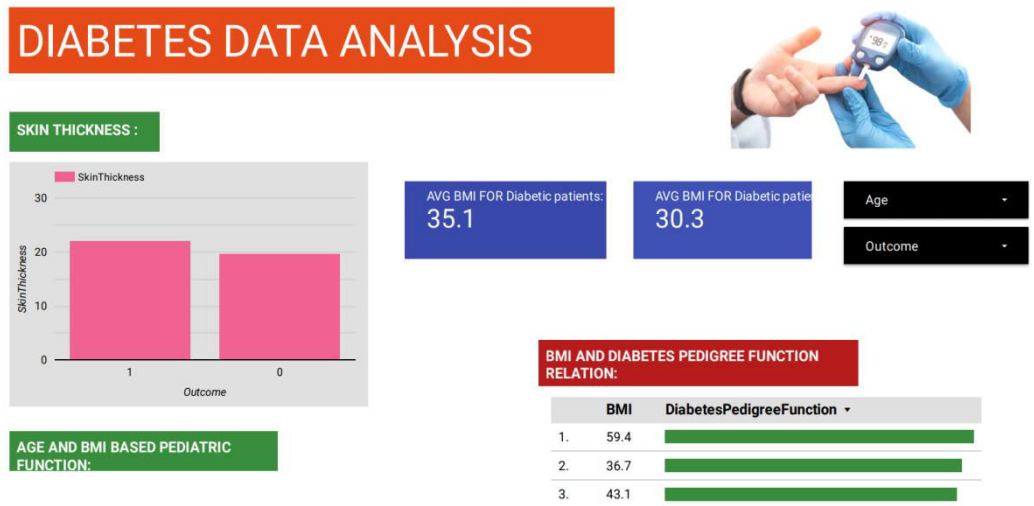


Fig 3: Data Analysis done on Data Studio

III. FUTURE SCOPE & DISCUSSION

Proposed system uses “KNN algorithm” to find the diabetes disease, in data science we have many algorithms for classification such as Naive Bayes, SVM, Decision Tree, ID3 etc… in future we can add more algorithms to find outputs and algorithms can be compared to find the efficient algorithm. We can add visitor query module, where visitors can post queries to administrator and admin can send reply to those queries. We can add treatment module, where doctors upload treatment details for patients and patient can view those treatment details.

The prediction of diabetes is one the of great importance in today scenario, and concerning with its severe complications. Due to the biggest reason for the death in worldwide is diabetes. The System model is mainly focus to identification of diabetes using some of the parameters. System is useful to physicians to predict the diabetes in initial dais. So, that conventional treatments and solutions may be given to the patients. System used some of the techniques like ML for the prediction, so that to get the more precise results. There have been fortune of investigation on the diabetes imprint. Building diabetes disease prediction system is useful for hospitals and doctors. System predicts disease at early stages, so doctors can treat patients in a better way. Proposed model is the real time application in which is meant for multiple hospitals and predicts disease in less time. As we use machine learning algorithms for disease prediction, we will get more accurate and efficient results.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905064 |

IV.CONCLUSION

Control of diabetes was significantly associated with social identity, the belief that diabetes is controlled by medication, self-monitoring blood glucose levels, and goal self-efficacy and lifestyle goal. Support group membership did improve goal behaviours, particularly medication goals. Social identity also influenced goal setting for lifestyle goals. Illness identity impacted lifestyle goal self-efficacy and medication taking goal behaviours. These findings have implications for the assessment of identity in the provision of healthcare services and the use of support groups in the goal setting process for the self-management of type 2 diabetes. From a patient care perspective, these findings suggest that online support groups have a positive impact on goal setting and achievement for those members that identify with the group and the objectives of the group. Health care providers that use goal setting to achieve clinical control of type 2 diabetes might find such groups useful tools to assist in setting and achieving goals. Support provided by such groups can also be helpful in educating the patient on methods of treating diabetes and improve the communication between the provider and the patient. On-going support is an essential component of mastering goals and healthcare providers should advocate this type of support in conjunction with regular check-ups and in addition to support from family, friends and care providers.

REFERENCES

1.Zimmet P, Alberti KGMM, Shaw J. Global and societal implications of the diabetes epidemic. Nature 2001; 414: 782-787.
2. Centers for Disease Control and Prevention. Take Charge of Your Diabetes. 4th edition. Atlanta: U.S. Department of Health and Human Services, 2007.
3. National Diabetes Fact Sheet 2007. Centers for Disease Control and Prevention, National Institutes of Health and American Diabetes Association.
4. Boltri JM, Okosun IS, Davis-Smith M, Vogel RL. HemoglobinA1c levels in diagnosed and undiagnosed Black, Hispanic and White persons with diabetes: from NHANES 1999-2000. Ethnicity&Disease 2005; 15: 562-567.
5. Weller Sc, Baer RD, Pachter LM, Trotter RT, Glazer M, Garcia de Alba Garcia JE, Klein RE. Latino beliefs about Diabetes.Diabetes Care 1999; 22: 722- 728.
6. Dallo FJ. James SA. Acculturation and Blood Pressure in a Communitybased Sample of Chaldean-American Women. J Immigrant Health 2000; 2(3): 145-153.
7. Dallo FJ, Borrell LN. Self-reported Diabetes and Hypertension Among ArabAmericans in the United States. Ethn& Dis 2006; 16: 699-705.
8. Jaber LA, Brown MB, Hammad A, Nowak SN, Zhu Q Gharfoor A, Herman WH. Epidemiology of Diabetes Among Arab Americans. Diabetes Care 2003; 26(2): 308-313. 9. Kridli SA, Fakhouri H. Health needs of Arab American women in Michigan. Michigan Nurse 2003; 76(6): 9-14.
10. Cramer JA. A systematic review of adherence with medications for diabetes. Diabetes Care 2004;27:1218-24.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details