

Teachers' Exploration of Data Science Subject: For the Evaluation of Students' Academic Performance

Ms. Thatapudi. Sahana¹, Dr. K. Ramesh², K. Balraj³, Dr. S. Naga Padma⁴,
Dr. K. Rama Krishna Reddy⁵, Dr. Sabitha Francis⁶

¹Professor, Department of CSE & IT, Garden City University, Bengaluru, India.

²Anurag Engineering College, Kodad, Telangana, India.

³Assistant Professor of English, JNTUH, Hyderabad, India.

⁴Associate Professor, CMR Engineering College, Hyderabad, India.

Assistant Professor, Department of English, B. V. Raju Institution of Technology, Narsapur, India.

⁶corresponding author, Professor, Department of English, Malla Reddy Engineering College, Secunderabad, Hyderabad, India.

KEYWORDS

Data science,
English, language,
interdisciplinary,
statistics, Data
Analytics

ABSTRACT:

The paper focuses on the need for knowledge of Computer Science, Statistics, and Data Science for English teachers in particular and other subject teachers at large. Academic teachers are good at expressing in English both written and spoken. But in the situations of numerical comparisons like which numerical number of the student is the outlier in the class, are grey areas for them. Hence this paper helps teachers to get an awareness of incorporating the concepts of Google Collab, statistics, and how to do comparative study using central tendency in Data Science in their English/other subjects classrooms

1. Introduction

The focal point of the paper will be if the academic teacher learns Data Science, how her/his teaching skills improve. The need of the hour is for the teacher who may have a reflexive mood of her teaching, and interpret her teaching to improvise/update her skills in the digital era. Further, she/he can get practical knowledge of how to interpret students' results so the teacher can become an academic adviser to enhance their language/subject study skills.

2. Objectives

- Connecting English/ other subject teacher profession to Data Science subject
- Help with Data Science for English/ other subject teachers
- Up skilling of English teachers

3. Research Questions

1. How is Data Science subject helpful for English teachers?
2. When can we connect English to Data Science?
3. Does upgrading of knowledge take place for English teachers with Data Science subjects?

4. Literary Background

In "Student -Permutator: Predicting Students' Academic Performance at Secondary and Intermediate level Using Machine Learning" Shah Hussain and Muhammad Qasimkhan advocate that, "As the data Science involves developing techniques of storing, and analyzing data to extract useful information efficiently".

In "Data Science for Analyzing and Improving Educational Processes" Shadi Aljawarneh and Juan A. Lara opine that The application of DS in the field of education may result of great interest for involved Stakeholders (students, instructors, and institutions....) since the extracted knowledge from educational data would be useful to deal with educational problems such as students ' improvement performance".

In “Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review” Abdallah Namoun , and Abdullah Alshanqiti explicitly present that “ (1) the forms in which the learning outcomes are predicted, (2) the predictive analytics models developed to forecast student learning, and (3) the dominant factors impacting student outcomes” by using data Analytics which is a part of Data Science.

5.Scope of the Research

As the researchers are dealing with the first-year B. Tech classes, the data is taken based on their academic performance. The students are also from the same institution where the researcher(corresponding author) is working.

6.Limitations of the Research

The B.Tech students of other domains of other institutions are not taken into consideration.

7.Research Design

The critical elements of research design are the study setting; measurement; operation of the research and research approach.

Operation of the Research: It is through conducting the Mid-I and Mid-II exams in a semester. And considering the final results for 40 marks.

7.1 Study Setting

The study setting is one semester of the I year B. Tech students.

The research is carried forward among students of I year B. Tech of Computer Science. Their final marks are taken as a sample study.

7.2 Measurement

The student's academic performance in the continuous evaluation is taken for the measurement to denote students' progress. The Mid-I syllabus is taught for 8 weeks and the Mid-II syllabus is taught for another 8 weeks. Their final marks for 40 are taken into consideration. It is done through Descriptive Statistics. Mean, Mode, Median, Standard Deviation (STD), and variance are found by considering the final marks.

7.3 Operationalization

The elements which are abstract or concrete concepts into measurable observations through some operations. Although some concepts like height or age are easily measured, others like spirituality or anxiety are not.

There are some students in the sample who express that they can give self-rating scores. However, some of the students expressed anxiety and agitated moods because of some long-term illness. Some show physical anxiety symptoms in social situations.

7.4 Research approach

- Inductive approach
- Deductive Approach
- Abductive Approach

7.4.1 Inductive approach

It's a bottom-up approach. It begins with the researchers collecting the data that is relevant to the research study. Post-data collection, a researcher will analyze the data.

It makes a generalization from specific observations and facts, while deductive reasoning uses available information, knowledge, or facts to construct a valid conclusion.

It helps for patterns in the data to develop a theory that could explain the patterns.

7.4.2 Deductive Approach

This approach is developing a hypothesis or hypotheses based on existing theory and then designing a research strategy. It is a top-down approach.

7.4.3 Abductive Approach

Inference is the best example of an abductive approach. The teacher must know whether the learning of the subject has taken place or not at the end of the teaching.

7.4.4 System Approach

The behavior of one component of a system affects the performance of other components due to interconnection and hence the system is studied as a whole. So it is called system approach.

8. Methodology

The methodology is a System Approach or scientific approach

If the data are small in size, the researchers can do it manually. But the numbers are many or huge, by applying the statistics, the researcher can easily calculate the data with the help of formulae in the computer program.

8.1 Benefits: Using mean, median, and Mode (in Descriptive Statistics), the data can be organized, analyzed, interpreted, and can be presented of data. It helps in drawing meaningful conclusions from data, using both descriptive and inferential methods.

8.2 English for Specific Purpose: English teachers understanding the technical subject can significantly help them. Particularly enhancing their ability to teach English in a particular technical domain.

9. Discussion

An English teacher with the students' academically performed data uses Descriptive Statistics to display the data. With the display, the viewers see the data distribution for better understanding. And sometimes they can also interact with the right visualization. Hence it would be possible everyone to be on the same page, regardless of their level of expertise.

Exploratory Data Analysis (EDA):

- Visualizing data through graphs and charts to identify trends and patterns
- Calculating descriptive statistics like mean, median, and Standard Deviation to understand and variability

10. Results and Interpretation

```
import pandas as pd
df = pd.read_excel(r"C:\Users\Pinaki\Downloads\2024-2025-CSE-English_for_Skill_Enhancement_Section-A.xlsx", )
df.head()
```

S.No	Roll No	Marks
0	1	25
1	2	39
2	3	30
3	4	33
4	5	23

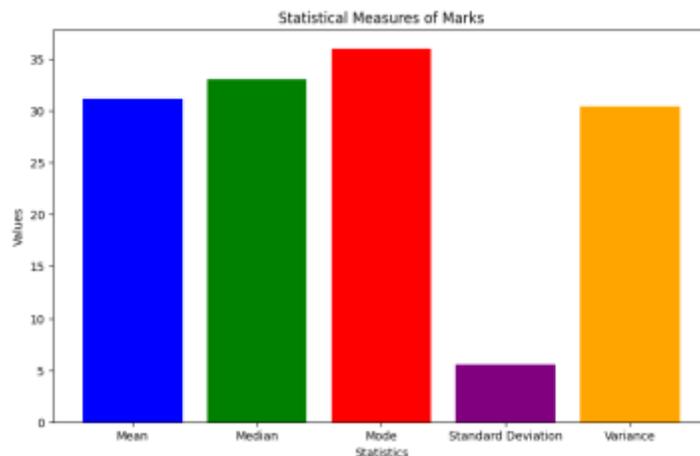
```
mean_marks = df['Marks'].mean()
median_marks = df['Marks'].median()
std_marks = df['Marks'].std()
```

```
var_marks = df['Marks'].var()
mode_marks = df["Marks"].mode()

print(f"Mean: {mean_marks}")
print(f"Median: {median_marks}")
print(f"Standard Deviation: {std_marks}")
print(f"Variance: {var_marks}")
print(f"Mode_marks: {mode_marks[0]}")
Mean: 31.16923076923077
Median: 33.0
Standard Deviation: 5.515798986687102
Variance: 30.424038461538462
Mode marks: 36
```

So, with the above data, the mean which is 31.169 denotes the average of the marks data. The average is 31.169 means the overall performance of the students is good (76%). And median:33 denotes the middle value of the set of numbers in the class. It means 50.7% of the class got desirable marks. Also, the mode: 36 denotes the value that appears frequently among the data.

```
import matplotlib.pyplot as plt
# Data for plotting
stats = ['Mean', 'Median', 'Mode', 'Standard Deviation', 'Variance']
values = [mean_marks, median_marks, mode_marks[0], std_marks, var_marks]
# Create a bar plot
plt.figure(figsize=(10, 6))
plt.bar(stats, values, color=['blue', 'green', 'red', 'purple', 'orange'])
plt.xlabel('Statistics')
plt.ylabel('Values')
plt.title('Statistical Measures of Marks')
plt.show()
```



11. Advantages of Using Data Science to evaluate the student's academic performance comprehensively

Applications of Data Science in Student Evaluation:

- Early Intervention:

Any academic teacher identifies students struggling early in the course, then they can come up with remedial courses for the targeted students thereby making provision to prevent academic decline.

- Personalized Learning:

Moderate students are comfortable with peer learning whereas a few students require special methods like tailored instructions based on individual student needs and learning styles.

- Continual Assessment Design:

The assessment method may have Optimized assessments by analyzing item difficulty. And specifically, whether the question has the qualities like concise, open-ended, and purposeful.

- Curriculum Improvement:

Redesigning the curriculum for the benefit of future studies and also identifying areas where students consistently struggle to cope. Hence the curriculum would be refined.

- Feedback Mechanism:

Providing more detailed and constructive feedback to students based on data analysis

12. Conclusion

12.1 Demographic Experiment on the Sample Group

14.1 After observing the class for approximately 16 weeks, the following observations are made: Some of the reasons for student's performance can be:

- Limited learning opportunity was found only among 10%;
- Students reported that “Transition in their study curriculum” happened only for 20% students;
- Among 65 students “Anxiety & stress” and stress are found among 10% of the students;
- Some of the students like 15% are Interest in different fields;
- As the researchers did not go into the family matters “Disordered home environment” is not known;
- It is noticeably found that “The basis is not as strong it should be” among the students;
- As Intermediate course in both Telugu states is different, the students found that “Complexity in study material”
- They happily expressed that there is absolutely no “Unsuitable methodology” from the faculty.

12.2 The researchers' observation of Achievers' academic performance, because of 1.5% of students are strong-willed to learn to the best of their abilities;

2. 15% have a positive attitude, though they are lagging in academics;

3. 15% of students have a love for learning, so they are meticulous in their learning behavior in the classroom;

4. Teachers felt that it was great to have them in researchers' class;

5. 10% of students are fantastic students who are always eager to learn new things;

6. For a committed teacher, 10% of students are treated to them;

7. 15% of students are quick learners who have a zeal learn; and

8. 30% of students are determined students to score higher credits.

Web -References

<https://link.springer.com/article/10.1007/s40745-021-00341-0>

<https://link.springer.com/article/10.1007/s12528-021-09299-7>

https://www.researchgate.net/publication/340890342_Exploring_Student_Academic_Performance_Using_Data_Mining_Tools

<https://www.tandfonline.com/doi/full/10.1080/26939169.2023.2208185>

<https://www.mdpi.com/2076-3417/11/1/237>

Print textbooks

Hira. DS, System Simulation. 2008. New Delhi: S. Chand & Company Ltd.