DATA MANAGEMENT PROJECT REPORT

(Project Semester August-December 2022)

on

(STUDENT PERFORMANCE IN EXAM DASHBOARD)

Submitted by

Siddharth Rathore (12100767)

Program Name: B. Tech CSE

Section – KM100

Under the Guidance of

Ms. Veerpal Kaur (25909)

School of Computer Science & Engineering Lovely Professional University Phagwara, Punjab.

CERTIFICATE

This is to certify that Siddharth Rathore bearing Registration no. 12100767 has completed INT217 project titled, **"Student Performance in Exam Dashboard"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort, and study.

Signature and Name of the Supervisor – Ms. Veerpal Kaur
Designation of the Supervisor - Professor
School of Computer Science & Engineering
Lovely Professional University
Phagwara, Punjab.

Date: 5th November 2022

DECLARATION

I hereby declare that I have completed my project of Student Performance in Exam Dashboard from 5th October 2022 to 5th November 2022 under the guidance of Ms. Veerpal Kaur. I have declared that I have worked with full dedication during these one month and my learning outcomes fulfill the requirements of project for the award of degree of B. Tech CSE, Lovely Professional University, Phagwara.

Name of Student: Siddharth Rathore

Registration no: 12100767

Date: 5th November 2022

ACKNOWLEDGEMENT

A few typewritten words of thanks cannot really express the sincerity of my gratitude. But I am still trying to put into words my gratefulness towards all who have helped and encouraged me in carrying out this Project. This project of mine bears the imprint of many people who have an important impact on my thinking, behavior, and acts during the technology learning and project making.

First, I would like to take this opportunity to thank the LOVELY PROFESSIONAL UNIVERSITY for having INT217 course as a part of the B. Tech CSE degree. The accomplishment of this project otherwise would have been painstaking endeavor, for lack of staunch and sincere support of the School of Computer Science and Engineering, LPU. The incessant and undeterred succors extended from the mentor of Lovely Professional University. They explained the course to the great extent. If this goes unnoticed and unacknowledged it would be selfishness.

Many people have influenced the shape and content of this Project, and many supported me throughout. I express my sincere gratitude to Ms. Veerpal Kaur, who was available for help whenever I required, their guidance, gentle persuasion and active support has made it possible to complete this Project.

In the end, I can say only this much that “ALL ARE NOT MENTIONED BUT NONE IS FORGOTTEN”

Last but not the least I would like to thank GOD, who continues to look after us despite all my flaws.

TABLE OF CONTENTS

INTRODUCTION

In this course of Introduction to Data Management which is part of Data Analysis where we do manipulation of data and make a meaningful data from the raw dataset.

Data Analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense, and recap, and evaluate data.

This project report consists of raw data of Students Performance in Exams. The dataset is in raw form, and I have used my excel skill to make it meaningful in a good manner.

In this project I have taken five main problem statement for the analysis which gives meaning to the dataset so that anyone can analyze what short of problems students are facing or what kind of resources are better for them, also checking the performance of the students through it.

This Project, I have worked on Student Exam Dataset in which pivot table and power pivot has been used. I created Dashboard, which is fully active with the help of slicer, for this project I have used power pivot to manage the dataset and for preparation of the dashboard. ETL process has been followed in this Project. A PivotTable is an interactive way to quickly summarize large amounts of data. You can use a PivotTable to analyze numerical data in detail and answer unanticipated questions about your data. A PivotTable is especially designed for: Querying large amounts of data in many user-friendly ways. Power Pivot is an Excel feature that enables the import, manipulation, and analysis of big data without loss of speed/functionality. Power Pivot tables are pivot tables that that allow the user to mix data from different tables, affording them powerful filter chaining when working on multiple tables. With Power Query (known as Get & Transform in Excel), you can import or connect to external data, and then shape that data, for example remove a column, change a data type, or merge tables, in ways that meet your needs. Then, you can load your query into Excel to create charts and reports. Analyze Data in Excel empowers you to understand your data through natural language queries that allow you to ask questions about your data without having to write complicated formulas. In addition, Analyze Data provides high-level visual summaries, trends, and patterns. The term data set refers to a file that contains one or more records. The record is the basic unit of information used by a program running on z/OS. Any named group of records is called a data set. Data visualization tools are software applications that render information in a visual format such as a graph, chart, or heat map for data analysis purposes. Such tools make it easier to understand and work with massive amounts of data.

We have also followed ETL process. This ETL process is the main lead which is used to give meaning to a raw dataset.

OBJECTIVES/SCOPE OF THE ANALYSIS

The five main objectives that I've worked on are as follows:

1. Comparison of Math's and Writing score Gender wise (Students).
   By this we can get the data of which gender is doing better. It is going to help in growing categorically gender wise.
2. Comparison of Reading and Writing score Gender wise (Student).
   By this we can get the data of which gender is doing better. It is going to help in growing categorically gender wise.
3. Completion of test preparation course of students Gender wise.
   By this we can get the data of which gender is more serious about the course. It is going to help in finding which student is lacking in the study.
4. Students Gender wise Lunch System.
   By this we can get the data of which gender is getting which kind of lunch system.
5. Students Gender wise Parental Level of Education.
   By this we can get the data of gender wise parental education quality.

SOURCE OF DATASET

Kaggle is an online community platform for data scientists and machine learning enthusiasts. Kaggle allows users to collaborate with other users, find and publish datasets, use GPU integrated notebooks, and compete with other data scientists to solve data science challenges.

The term data set refers to a file that contains one or more records. The record is the basic unit of information used by a program running on OS. Any named group of records is called a data set.

Dataset Link: https://www.kaggle.com/datasets/whenamancodes/students-performance-in-exams

Dashboard Link: https://github.com/s1dbugs/Analysis-on-Student-Performance-in-Exam

ETL PROCESS

The 5 steps of the ETL process are: extract, clean, transform, load, and analyze. Of the 5, extract, transform, and load are the most important process steps.

- Extract: Retrieves raw data from an unstructured data pool and migrates it into a temporary, staging data repository. For this I have used kaggle.com, many people over there are posting unlimited number of datasets. We can get any raw dataset from anywhere.
- Clean: Cleans data extracted from an unstructured data pool, ensuring the quality of the data prior to transformation. For that I have used power pivot to clean the data which I got from kaggle.com that is Student performance in exam. By the term cleaning I means checking the Duplicate, Blank space, and unorganized form of data. Cleaning of dataset is must for every analyst because after cleaning data will be easier to manipulate.
- Transform: Structures and converts the data to match the correct target source. Its easy to transform data when its cleaned properly. It shows each field which is going to load is proper or not.
- Load: Loads the structured data into a data warehouse so it can be properly analyzed and used. (For transform and load data I use one menu of power pivot i.e., Manage.)
- Analyze: Big data analysis is processed within the warehouse, enabling the business to gain insight from the correctly configured data. For that I use pivot table, power pivot and power query editor.

ANALYSIS ON DATASET

- **Introduction**

  In this Project, I have worked on Student Exam Dataset in which pivot table and power pivot has been used. I created Dashboard, which is fully active with the help of slicer, I have taken five main problem statement for the analysis.

  For this project I have used power pivot to manage the dataset and for preparation of the dashboard. I have used tools such as pivot chart which is pie chart, bar chart & line chart for better understanding of the dashboard, I have used two slicer, gender and race/ethnicity which helps to show exact values in the chart, also I've used page layout for better representation to the dashboard, I have done gridline editing to representing the background, for title of the sheet I have used merge tool. ETL process has been followed in this Project. Everything in the project has gone through the ETL process. The problem statement is useful in such way to improve the student's performance that will help them to improve their score. ETL means Extract Transform and Load which comes under data representation.

- **General Description**

  In this project I have basically analysed that,

  - Male students are good in maths and Female students are good in writing.
  - Maximum male students have completed test preparation course than female students.
  - Female students take standard lunch system more than male students.
  - Parental level of education is higher in male students than female students.
  - Students are better in every field but gender wise.
  - By their race/ethnicity and gender we can fetch exact score of the student.
  - Using of slicer make it easy to handle the dashboard.

- **Specific Requirements**

  - Pivot Table- A PivotTable is an interactive way to quickly summarize large amounts of data. You can use a PivotTable to analyse numerical data in detail and answer unanticipated questions about your data. A PivotTable is

especially designed for: Querying large amounts of data in many user-friendly ways.

- Power Pivot- Power Pivot is an Excel feature that enables the import, manipulation, and analysis of big data without loss of speed/functionality. Power Pivot tables are pivot tables that that allow the user to mix data from different tables, affording them powerful filter chaining when working on multiple tables.
- Power Query Editor- With Power Query (known as Get & Transform in Excel), you can import or connect to external data, and then shape that data, for example remove a column, change a data type, or merge tables, in ways that meet your needs. Then, you can load your query into Excel to create charts and reports.

- **Analysis Results**

  After the use of every analysis tool, I am getting the result that Female gender students is poor in some of the subjects, but also good in some other areas than Male gender students. This analysis is basically trying to find the student performance depending upon their parental level of education or subjects score comparison and lunch system. By this we can improvise the system resource failure in each gender. Every student is having their own strong subject

  Recommendation: By giving better valuable resources of maths subject we can improve the result of mathematics of female students simultaneously by giving more good resources we can improve the writing score of male students.

  Also, by the analysis I observed that because of parental education girls are lacking in scoring more numbers in the high weightage subject so we can also upgrade and update parents about their child which can help both the gender in improving the result.

- **Visualization**

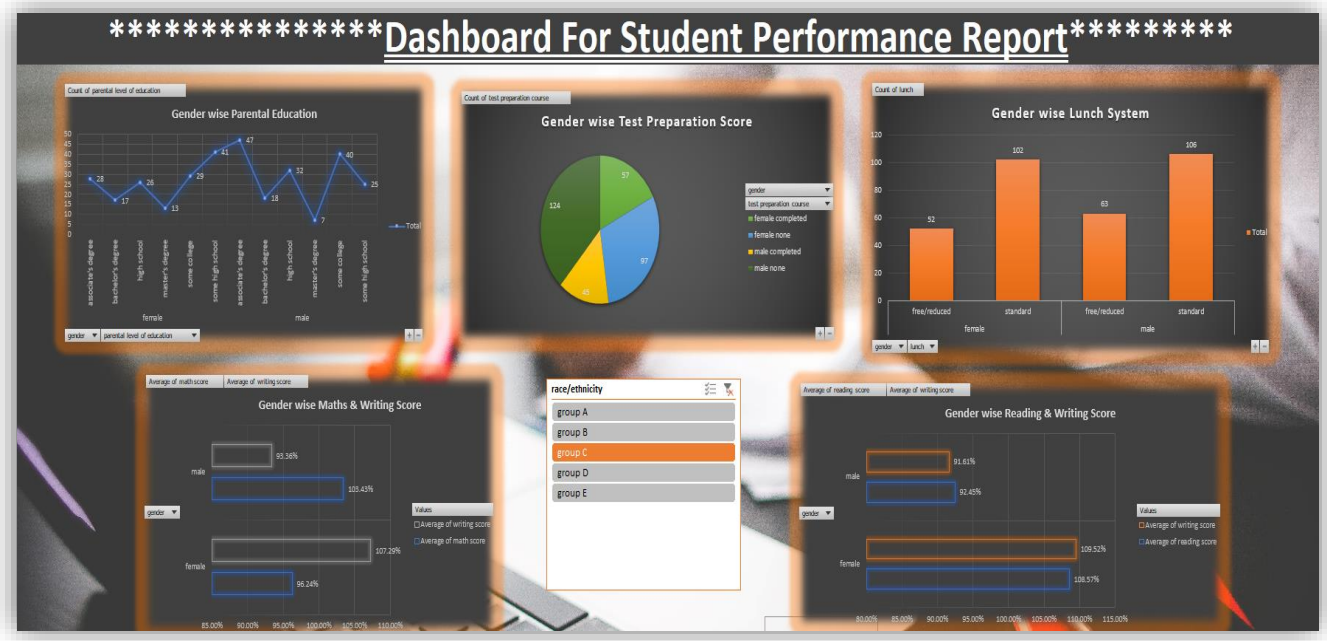

Fig. 5.1 Dashboard

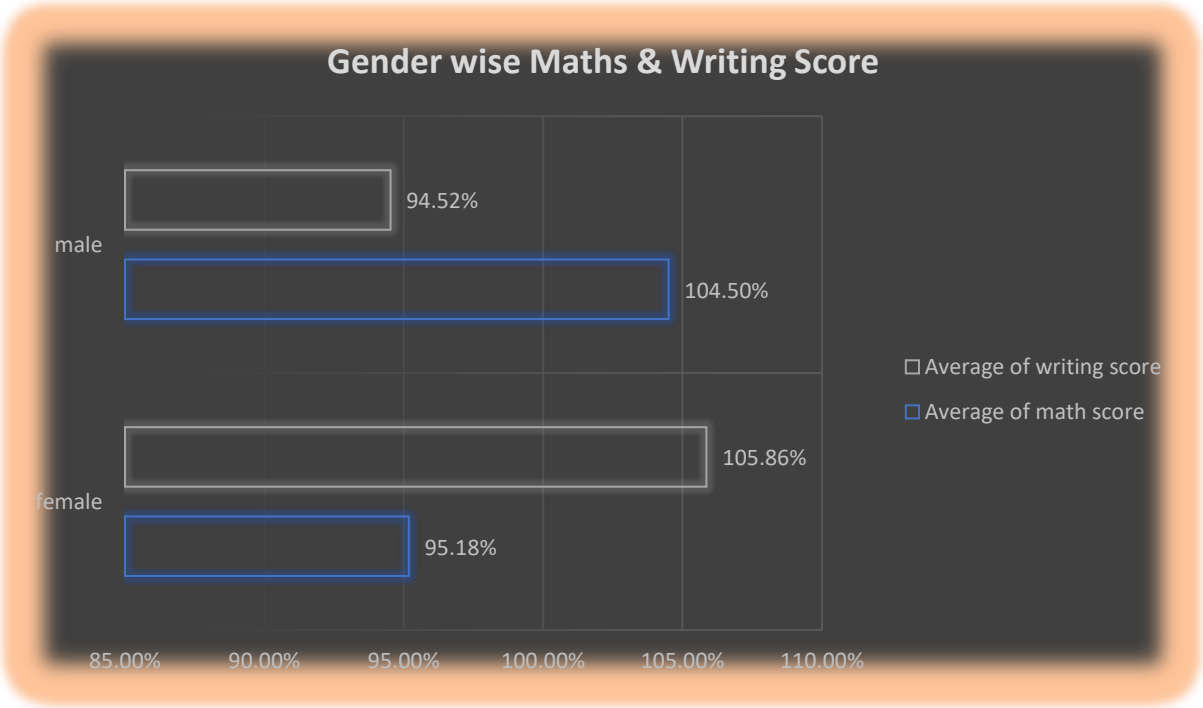

Fig 5.2

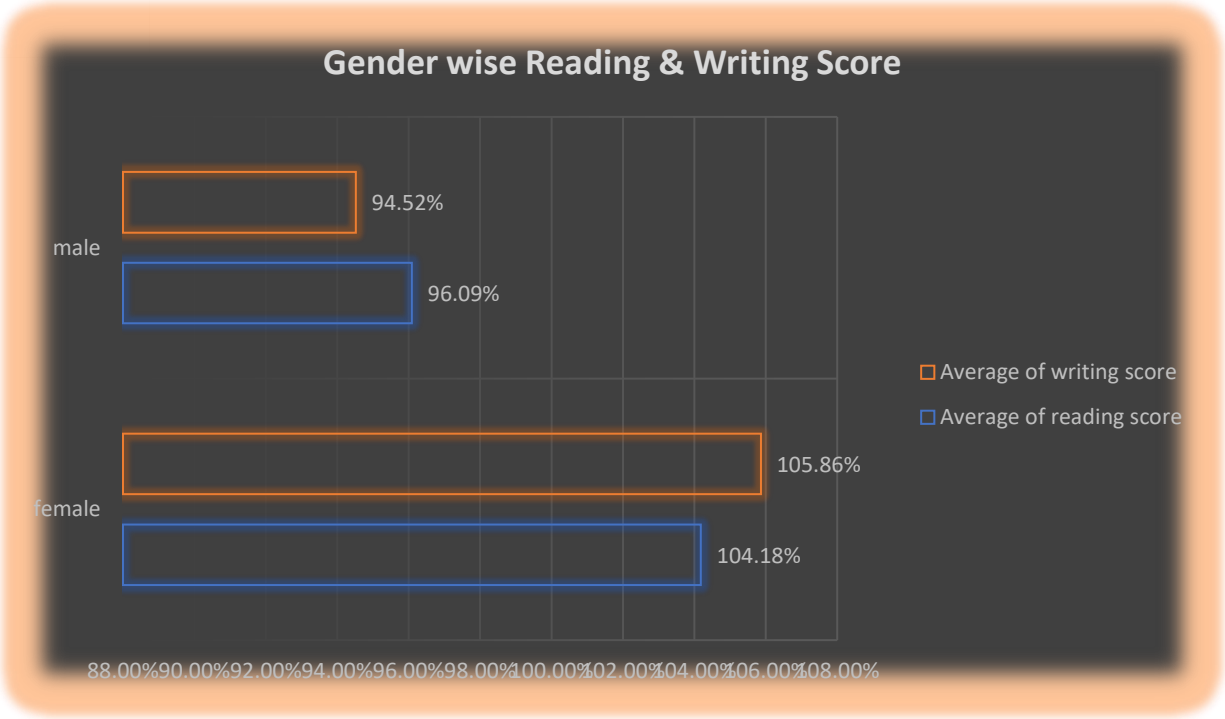
Gender wise Reading & Writing Score
male
female
94.52%
96.09%
105.86%
104.18%
Average of writing score
Average of reading score
88.00%
90.00%
92.00%
94.00%
96.00%
98.00%
100.00%
102.00%
104.00%
106.00%
108.00%

Fig 5.3

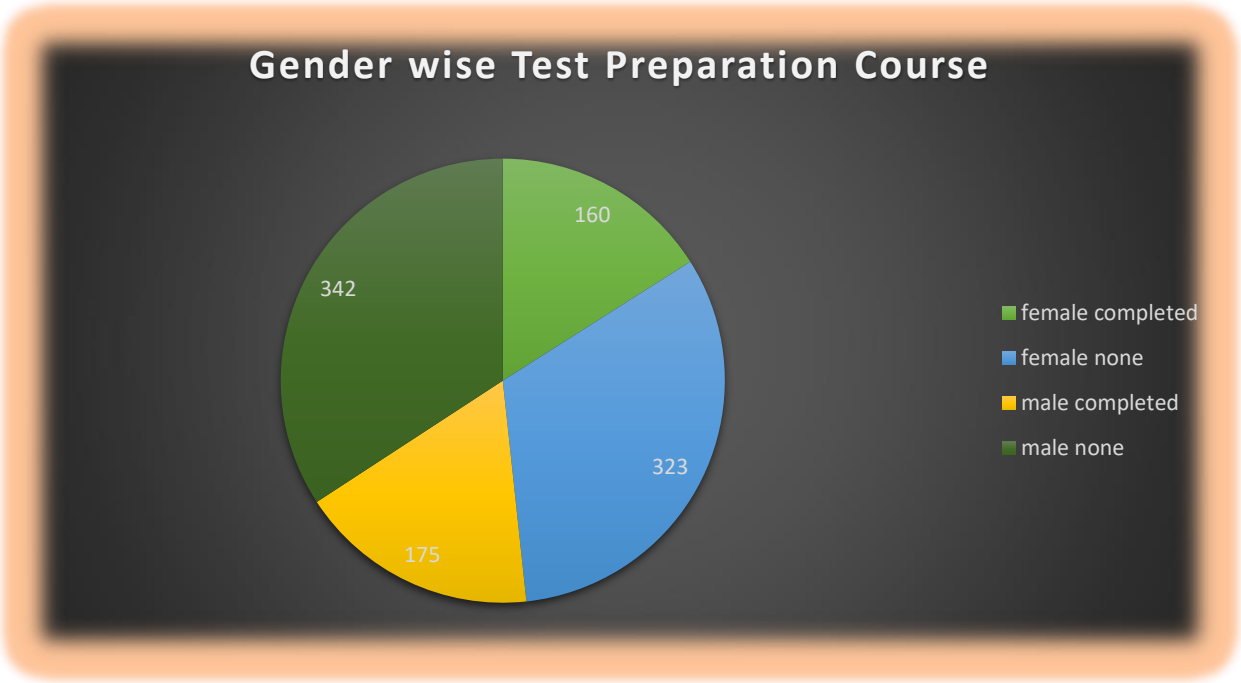
Gender wise Test Preparation Course
160
323
175
342
female completed
female none
male completed
male none

Fig 5.4

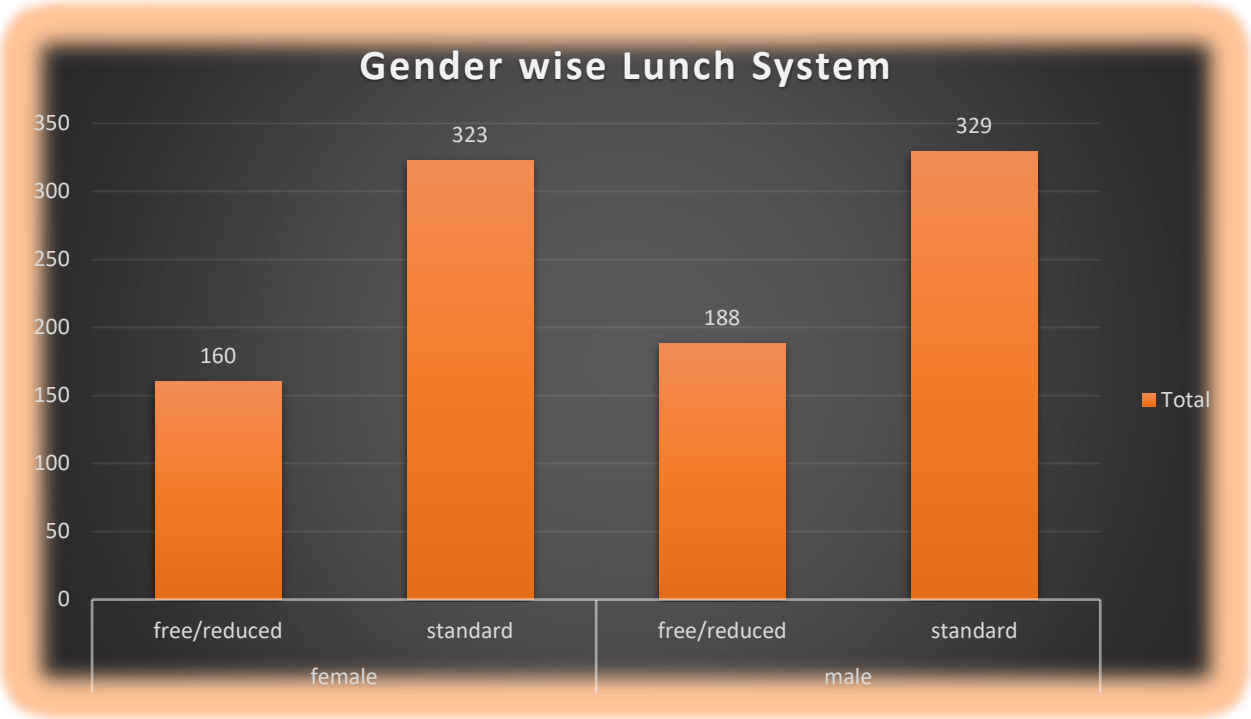
Gender wise Lunch System
350
300
250
200
150
100
50
0
160
323
188
329
free/reduced
standard
free/reduced
standard
female
male
Total

Fig 5.5

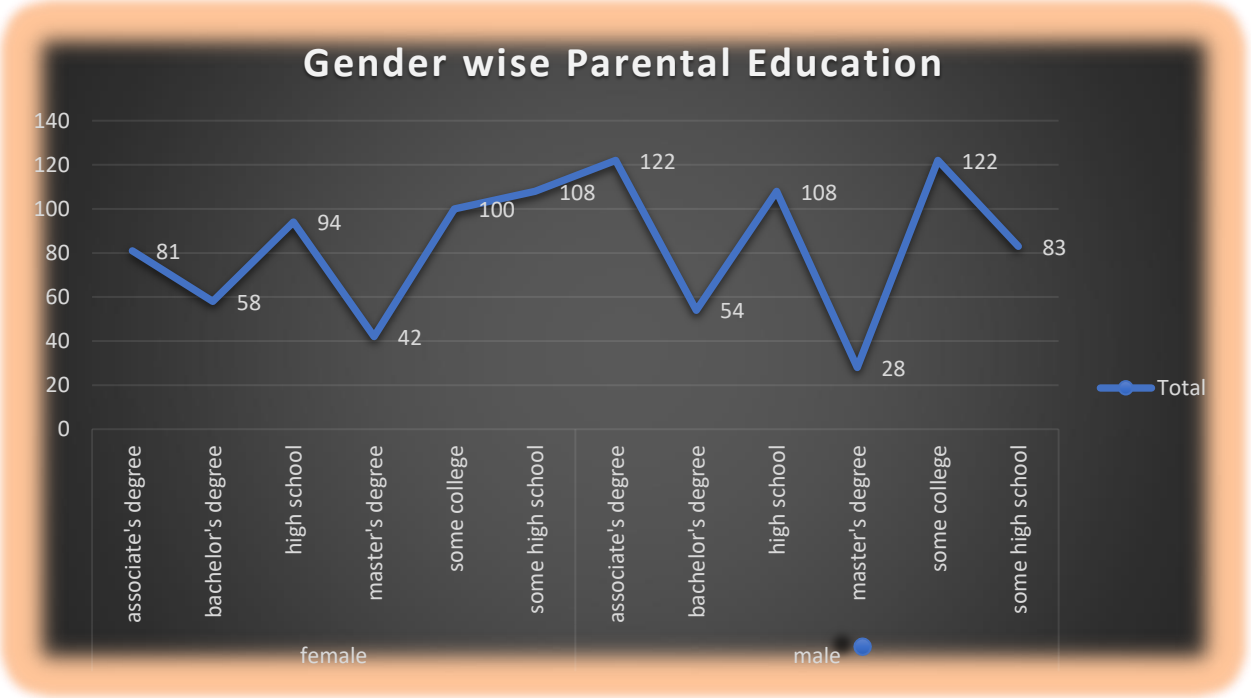
Gender wise Parental Education
140
120
100
80
60
40
20
0
81
58
94
42
100
108
122
54
108
28
122
83
associate's degree
bachelor's degree
high school
master's degree
some college
some high school
associate's degree
bachelor's degree
high school
master's degree
some college
some high school
female
male
Total

Fig 5.6

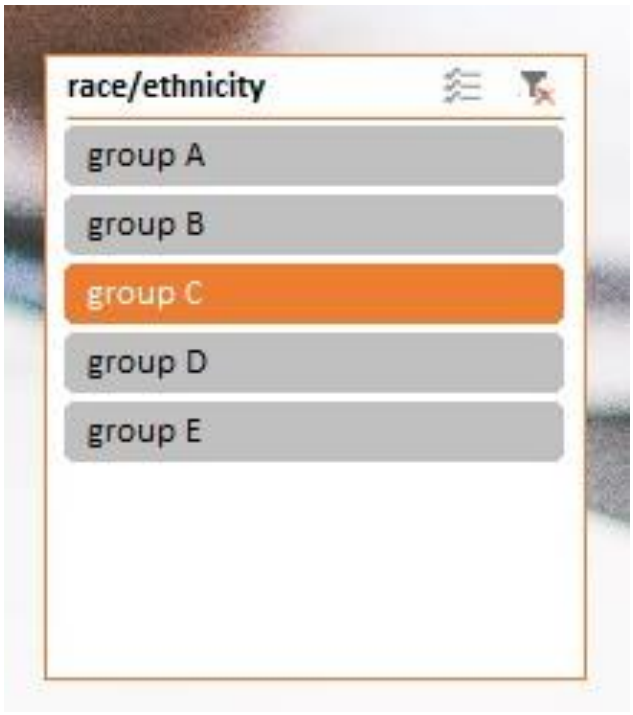

Fig 5.7 Slicer

LIST OF ANALYSIS WITH RESULTS

1. Comparison of Math's and Writing score Gender wise (Students).

   In this analysis by using bar chart, I've observed that the male students are good in math's & poor in writing and the female students are good at writing but poor in math's.

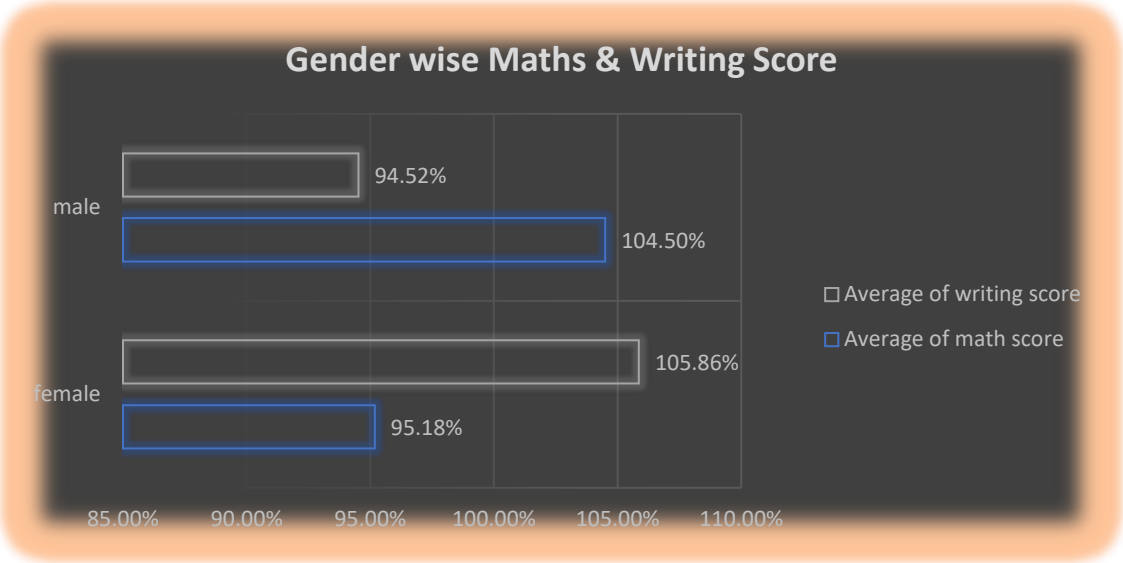

Fig. 1

2. Comparison of Reading and Writing score Gender wise (Student).

   In this analysis by using bar chart, I've observed that the female students are good in reading & writing but male students are average in reading & writing.

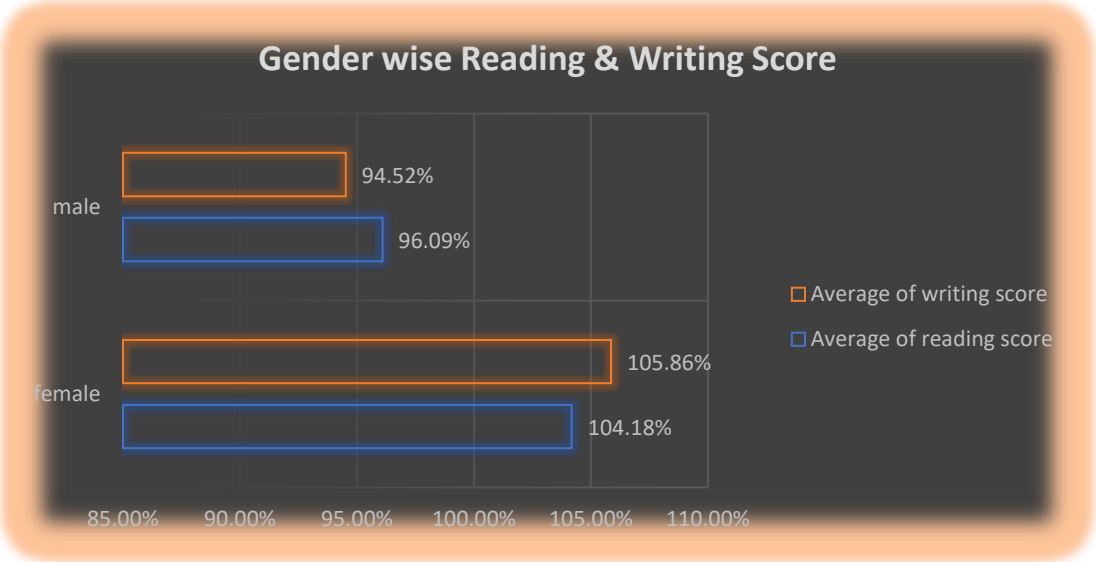

Fig. 2

3. Completion of test preparation course of students Gender wise.

   In this analysis by using pie chart, I've observed that female test preparation course is less completed than male and non-completion of test preparation course of female is higher than male.

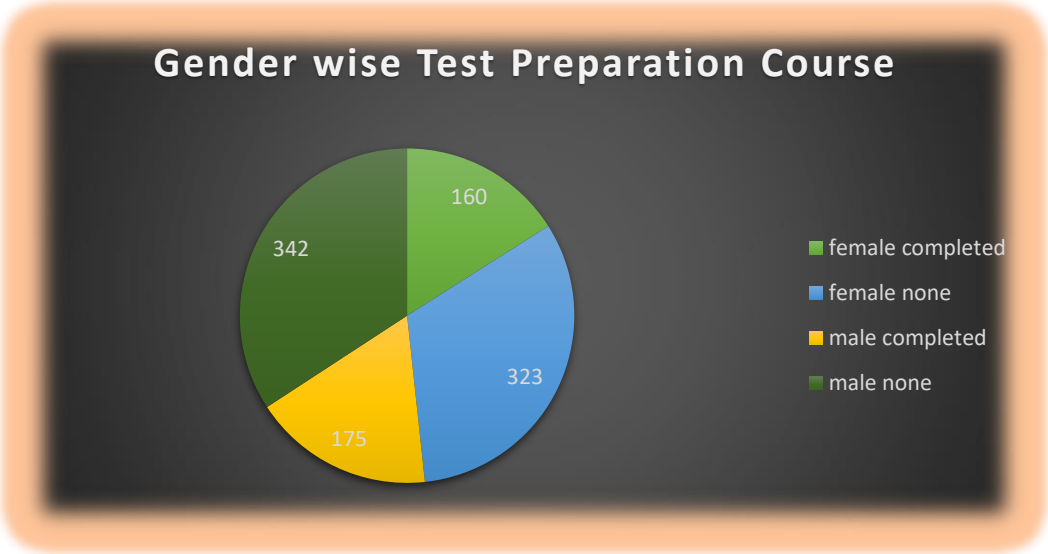

Fig. 3

4. Students Gender wise Lunch System.

   In this analysis by using column chart, I've observed that female students are using standard lunch system than male students, but male students are using free/reduced lunch system more than female students.

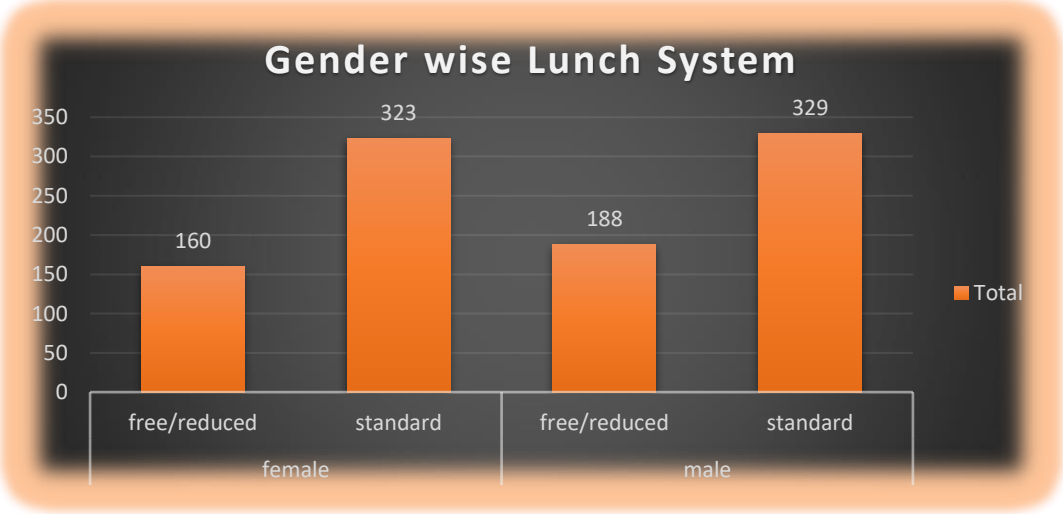

Fig. 4

5. Students Gender wise Parental Level of Education.

   In this analysis by using line chart, I've observed that max. female student parents' education level are some high school level, but male student parents education level are equal in associate degree and some high school level.

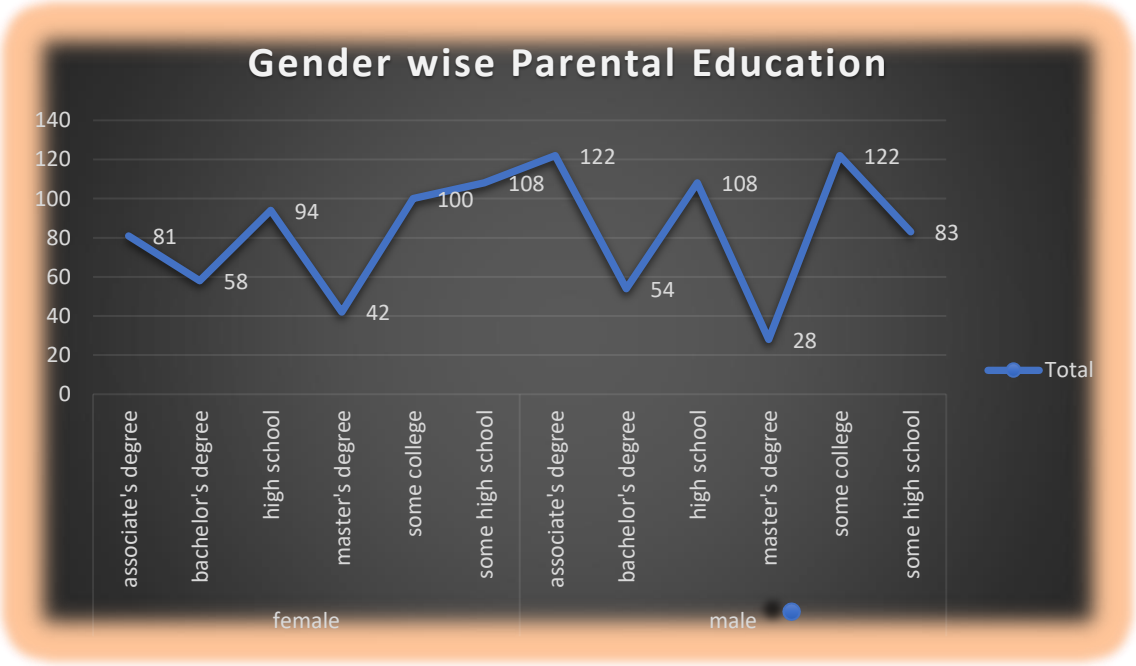

Fig. 5

REFERENCES

I was able to get all the data I required to create the dashboard, some help was taken from various Websites and YouTube for design inspiration for my project. I came up with my own ideas and tried several methods on my own to complete the project. However, we did use certain concepts from www.javatpoint.com, www.github.com, and www.geeksforgeeks.org, for specific dashboard components.

Geeks for geeks: https://www.geeksforgeeks.org/

Course Hero: https://www.coursehero.com/

GIT Link: https://github.com/s1dbugs/

Kaggle: https://www.kaggle.com/

BIBLIOGRAPHY

Tutorial's point -

https://www.tutorialspoint.com/excel_dashboards/excel_dashboards_tutorial.pdf

Dashboard Ideas-

https://analysistabs.com/project/management/dashboards/excel-template/

THE END